**Eigenvalue Problem of Symmetric Matrix**

In a vector space, if the application of an operator $O$ to a vector ${\bf x}$ results in another vector $\lambda {\bf x}$, where $\lambda$ is constant scalar:

\begin{displaymath}
O{\bf x}=\lambda {\bf x},
\end{displaymath}

then the scalar $\lambda$ is an eigenvalue of $T$ and vector ${\bf x}$ is the corresponding eigenvector or eigenfunctions of $O$, and the equation above is called the eigenequation of the operator $T$. The set of all eigenvalues of an operator is called the spectrum of the operator.

Note that if ${\bf x}$ is an eigenvector of operator $O$ then $c{\bf x}$ is also an eigenvector.

In a function space, the $n$th-order differential operator $D^n=d^n/dt^n$ is a linear operator with the following eigenequation:

\begin{displaymath}
D^n\phi(t)=D^n\; e^{st}=\frac{d^n}{dt^n}\; e^{st}=s^n\;e^{st}=\lambda \phi(t),
\end{displaymath}

where $s$ is a complex scalar. Here, the $\lambda=s^n$ is the eigenvalue and the complex exponential $\phi(t)=e^{st}$ is the corresponding eigenfunction. More generally, we can write an $N$th-order linear constant coefficient differential equation (LCCDE) as

\begin{displaymath}
\sum_{n=0}^N a_n\frac{d^n}{dt^n} y(t)=\left[\sum_{n=0}^N a_nD^n\right] y(t)=x(t),
\end{displaymath}

where $\sum_{n=0}^N a_nD^n$ is a linear operator that is applied to function $y(t)$, representing the response of a linear system to an input $x(t)$. Obviously, the same complex exponential $\phi(t)=e^{st}$ is also the eigenfunction corresponding to the eigenvalue $\lambda=\sum_{k=0}^n a_ks^k$ of this operator.

Perhaps the most well-known eigenvalue problem in physics is the Schrödinger equation, which describes a particle in terms of its energy and the de Broglie wave function. Specifically, for a 1-D stationary single particle system, we have

\begin{displaymath}
\hat{{\cal H}} \psi(x)
=\left[-\frac{\hbar^2}{2m} \frac{\partial^2}{\partial x^2}+V(x)\right] \psi(x)
={\cal E}\psi(x),
\end{displaymath}

where

\begin{displaymath}
\hat{{\cal H}}=-\frac{\hbar^2}{2m} \frac{\partial^2}{\partial x^2}+V(x)
\end{displaymath}

is the Hamiltonian operator, $\hbar$ is the Planck constant, $m$ and $V(x)$ are the mass and potential energy of the particle, respectively. ${\cal E}$ is the eigenvalue of $\hat{{\cal H}}$, representing the total energy of the particle, and the wave function $\psi(x)$ is the corresponding eigenfunction, also called eigenstate, representing probability amplitude of the particle; i.e., $\vert\psi(x)\vert^2$ is the probability for the particle to be found at position $x$.

An $N$ by $N$ full-rank matrix ${\bf A}$ is a linear operator and the associated eigenequation is

\begin{displaymath}
{\bf A}{\bf\phi}_n=\lambda_n {\bf\phi}_n \;\;\;\;\;\;\;n=1,\ldots,N,
\end{displaymath}

where $\lambda_n$ and ${\bf\phi}_n$ are the $n$th eigenvalue and the corresponding eigenvector of ${\bf A}$, respectively. This equation can also be written as

\begin{displaymath}
{\bf A}{\bf\phi}_n=\lambda_n {\bf I}{\bf\phi}_n \;\;\;\;\;\...
...e.,}
\;\;\;\;\;\;\;
({\bf A}-\lambda{\bf I}){\bf\phi}={\bf0}
\end{displaymath}

We see that an eigenvector ${\bf\phi}$ is a non-zero solution of this homogeneous equation system, and the eigenvalues are the roots of the Nth order function

\begin{displaymath}f(\lambda)=det({\bf A}-\lambda{\bf I})=0 \end{displaymath}

representing the condition for non-zero solutions to exist.

The $N$ independent eigenvectors can be considered as the column vectors of an orthogonal matrix ${\bf\Phi}$:

\begin{displaymath}{\bf\Phi}=[{\bf\phi}_1,\cdots,{\bf\phi}_N] \end{displaymath}

and the $N$ eigen-equations can be written in matrix form as:

\begin{displaymath}
{\bf A}{\bf\Phi}={\bf A}[{\bf\phi}_1,\cdots,{\bf\phi}_N]
=...
... 0&\cdots&\lambda_N\end{array}\right]
={\bf\Phi}{\bf\Lambda}
\end{displaymath}

i.e., the matrix ${\bf A}$ can be diagonalized by its eigenvector matrix, i.e., ${\bf A}$ is similar to its eigenvalue matrix ${\bf\Lambda}$:

\begin{displaymath}
{\bf\Phi}^{-1}{\bf A}{\bf\Phi}={\bf\Lambda}
\end{displaymath}

If the matrix ${\bf A}^T={\bf A}$ is real and symmetric, then its eigenvalues are real and eigenvectors are orthogonal to each other, i.e., ${\bf\Phi}$ is orthogonal

\begin{displaymath}
{\bf\Phi}^{-1}={\bf\Phi}^T,
\;\;\;\;\;\;\;\mbox{i.e.}\;\;\;\;\;\;\;\;
{\bf\Phi}{\bf\Phi}^T={\bf\Phi}^T{\bf\Phi}={\bf I}
\end{displaymath}

and can be considered as a rotation matrix, and we have

\begin{displaymath}{\bf\Phi}^{-1}{\bf A}{\bf\Phi}={\bf\Phi}^T{\bf A}{\bf\Phi}={\bf\Lambda} \end{displaymath}

Before discussing Jacobi's method for finding ${\bf\Phi}$ and ${\bf\Lambda}$, we first review the rotation in a 2-D space:

\begin{displaymath}
{\bf R}(\theta)=\left[\begin{array}{rr}\cos\theta & \sin\the...
...ght] =\left[\begin{array}{rr}c & s -s & c\end{array}\right]
\end{displaymath}

where $c=\cos\theta$ and $s=\sin\theta$. Any vector ${\bf x}=[x_1, x_2]^T$ is rotated by an angle of $\theta$ by ${\bf R}(\theta)$ to become:

\begin{displaymath}{\bf R}(\theta){\bf x}=\left[\begin{array}{rr}c & s -s & c\...
...egin{array}{r}cx_1+sx_2 -sx_1+cx_2\end{array}\right]={\bf y} \end{displaymath}

the row or column vectors the rotation matrix ${\bf R}$ are orthogonal to each other, i.e., ${\bf R}$ is orthogonal satisfying ${\bf R}^{-1}={\bf R}^T$. This is in general true for any N-D rotation matrix.

**Example:** Let $\theta=\pi/4$, $\cos(\theta)=\sin(\theta)=1/\sqrt{2}$, and ${\bf x}=[2,1]^T$, then we have

\begin{displaymath}{\bf y}={\bf R}{\bf x}
=\frac{1}{\sqrt{2}}\left[\begin{array}...
...frac{1}{\sqrt{2}}\left[\begin{array}{r}1 3\end{array}\right] \end{displaymath}

The same rotation can be carried out in a 3-D space around any of the three axes with the rotation matrices:

\begin{displaymath}{\bf R}_z=\left[\begin{array}{rrr}c&s&0 -s&c&0 0&0&1\end{...
...eft[\begin{array}{rrr}c&0&-s 0&1&0 s&0&c\end{array}\right] \end{displaymath}

In an n-D space, the rotation matrix takes the following form

|  |  |  |  |
| --- | --- | --- | --- |
| $\displaystyle {\bf R}(i,j,\theta)$ | $\textstyle =$ | $\displaystyle \left[\begin{array}{rrrrrrr} 1&\cdots&0&\cdots&0&\cdots&0\\ \... ...s&0&\cdots&1\end{array}\right]\begin{array}{c}  i    j    \end{array}$ |  |
|  |  | $\displaystyle \;\;\;\;\;\;\;\;\;\begin{array}{ccccccccc}&&i&&&&j&&\end{array}$ |  |

This is an identical matrix with four of its elements modified, $r_{ii}=r_{jj}=c=\cos\theta$ and $r_{ij}=-r_{ji}=s=\sin\theta$. When multiplied by this rotation matrix, any vector in the n-D space will be rotated around the direction of $i\times j$ by an angle of $\theta$. Note that the rotation matrix is always orthogonal, i.e., its columns (or rows) are orthogonal to each other.

Jacobi method finds the eigenvalues of a symmetric matrix ${\bf A}$ by iteratively rotating its row and column vectors by a rotation matrix ${\bf R}$ in such a way that all of the off-diagonal elements will eventually become zero, and the diagonal elements are the eigenvalues.

Applying a rotation matrix ${\bf R}(i,j,\theta)$ to a symmetric matrix ${\bf A}={\bf A}^T$ we get

\begin{displaymath}
{\bf A}'={\bf R}^T{\bf A}{\bf R}
=\left[\begin{array}{cccc...
...s\\
&\cdots&a'_{ni}&\cdots&a'_{nj}&\cdots&\end{array}\right]
\end{displaymath}

Note that ${\bf A}'^T={\bf A}'$ is also symmetric:

\begin{displaymath}{\bf A}'^T=({\bf R}^T{\bf A}{\bf R})^T={\bf R}^T{\bf A}^T{\bf R}={\bf A}' \end{displaymath}

and all rows and columns of ${\bf A}'$ are the same as those of ${\bf A}$, i.e., $a'_{kl}=a_{kl}$ for all $k,l\ne i,j$, except the ith and jth rows and columns, which are updated:

|  |  |  |  |
| --- | --- | --- | --- |
| $\displaystyle a'_{kl}$ | $\textstyle =$ | $\displaystyle a_{kl}\;\;\;\;\;\;\;\;\;\;\;\;\;(k,l\ne i,j)$ |  |
| $\displaystyle a'_{ik}$ | $\textstyle =$ | $\displaystyle a'_{ki}=c a_{ik}-s a_{jk}\;\;\;\;\;\;\;\;\;(k\ne i,\;\;k\ne j)$ |  |
| $\displaystyle a'_{jk}$ | $\textstyle =$ | $\displaystyle a'_{kj}=s a_{ik}+c a_{jk}\;\;\;\;\;\;\;\;\;(k\ne i,\;\;k\ne j)$ |  |
| $\displaystyle a'_{ij}$ | $\textstyle =$ | $\displaystyle a'_{ji}=(c^2-s^2)a_{ij}+cs(a_{ii}-a_{jj})$ |  |
| $\displaystyle a'_{ii}$ | $\textstyle =$ | $\displaystyle c^2a_{ii}-2cs a_{ij}+s^2a_{jj}$ |  |
| $\displaystyle a'_{jj}$ | $\textstyle =$ | $\displaystyle s^2a_{ii}+2cs a_{ij}+c^2a_{jj}$ |  |

As the rotation matrix ${\bf R}$ is an orthogonal matrix, it does not change the norm (length) of the row and column vectors of the matrix ${\bf A}$. If we can set an off-diagonal element to zero $a_{ij}=0$ by some rotation matrix ${\bf R}(i,j,\theta)$, then the values of the diagonal elements $a'_{ii}$ and $a'_{jj}$ will be increased. The Jacobi method is to repeatedly carry out such rotations so that eventually all off-diagonal elements of the matrix become zero, i.e, ${\bf A}$ is converted into a diagonal eigenvalue matrix ${\bf\Lambda}$, by a sequence of orthogonal rotation matrices whose product is the eigenvector matrix ${\bf\Phi}$.

Setting the off-diagonal element $a'_{ij}$ to zero:

\begin{displaymath}a'_{ij}=(c^2-s^2)a_{ij}+cs(a_{ii}-a_{jj})=0 \end{displaymath}

we get

\begin{displaymath}
\frac{a_{jj}-a_{ii}}{a_{ij}}=\frac{c^2-s^2}{cs}=\frac{1-(s/c)^2}{s/c}=\frac{1-t^2}{t}=2w
\end{displaymath}

where we have defined

\begin{displaymath}t=\tan\theta=\frac{s}{c}=\frac{\sin\theta}{\cos\theta},
\;\;\...
...c^2-s^2}{2cs}=\frac{\cos(2\theta)}{\sin(2\theta)}=cot(2\theta) \end{displaymath}

We can now solve the equation above

\begin{displaymath}1-t^2-2wt=0,\;\;\;\;\;\;\;\mbox{i.e.,}\;\;\;\;\;\;t^2+2wt-1=0 \end{displaymath}

to get

\begin{displaymath}t_{1,2}=-w\pm\sqrt{w^2+1} \end{displaymath}

We will always choose the root with the smaller absolute value:

\begin{displaymath}t=\left\{\begin{array}{ll} -w+\sqrt{w^2+1}&\;\;\;\;\;\;\;\;w>0\\
-w-\sqrt{w^2+1}&\;\;\;\;\;\;\;\;w<0\end{array}\right.
\end{displaymath}

Having found $t=\tan\theta$, we can find the rotation angle $\theta$ and further find $c$ and $s$ needed in ${\bf R}$:

\begin{displaymath}s=\sin\theta=\frac{\tan\theta}{\sqrt{1+\tan^2\theta}}=\frac{t...
...s\theta=\frac{1}{\sqrt{1+\tan^2\theta}}=\frac{1}{\sqrt{1+t^2}} \end{displaymath}

Other than $a'_{ij}=0$ which is set to zero by the rotation, the rest of elements in ${\bf A}'$ can be updated by the set of equations given above.

To minimize roundoff error, it is preferable to update the elements by adding a term to its old value. To do so, we first solve the equation above $a'_{ij}=0$ for $a_{ii}$ and $a_{jj}$ to get

\begin{displaymath}\left\{\begin{array}{l}
a_{ii}=a_{jj}-\frac{c^2-s^2}{sc}a_{ij...
... \\
a_{jj}=a_{ii}+\frac{c^2-s^2}{sc}a_{ij}
\end{array}\right. \end{displaymath}

Substituting these into the expressions for $a'_{jj}$ and $a'_{ii}$ respectively, we get

\begin{displaymath}\left\{\begin{array}{l}
a'_{jj}=a_{jj}+t a_{ij}\nonumber  a'_{ii}=a_{ii}-t a_{ij}\end{array}\right.\end{displaymath}

We then rewrite the update equations for $a'_{ik}$ and $a'_{jk}$ as:

\begin{displaymath}\left\{\begin{array}{l}
a'_{ik}=c a_{ik}-s a_{jk}=a_{ik}+(c-...
... =a_{jk}+s(a_{ik}-\tau a_{jk})
\nonumber
\end{array}\right.
\end{displaymath}

where

\begin{displaymath}\tau=\frac{1-c}{s}=\frac{s}{1+c}=\frac{1-\cos\theta}{\sin\theta}=\frac{\sin\theta}{1+\cos\theta}
=\tan(\theta/2) \end{displaymath}

If we set an off-diagonal element to zero $a_{ij}=0$ by ${\bf R}(i,j,\theta)$, with $c=\cos\theta$ and $s=\sin\theta$ defined above, then the values of the diagonal elements $a'_{ii}$ and $a'_{jj}$ of the resulting matrix ${\bf A}'={\bf R}^T{\bf A}{\bf R}$ will be increased. If we iteratively carry out such rotations to set the off-diagonal elements to zero one at a time

|  |  |  |  |
| --- | --- | --- | --- |
|  |  | $\displaystyle {\bf A}''={\bf R}^T_2{\bf A}'{\bf R}_2={\bf R}^T_2{\bf R}^T_1{\bf A}{\bf R}_1{\bf R}_2$ |  |
|  |  | $\displaystyle \cdots \cdots \cdots \cdots \cdots \cdots$ |  |
|  |  | $\displaystyle {\bf A}^{(k)}={\bf R}^T_k\cdots{\bf R}^T_1{\bf A}{\bf R}_1\cdots{\bf R}_k$ |  |

until eventually the resulting matrix ${\bf A}^{(k)}={\bf\Lambda}$ becomes diagonal containing the eigenvalues of ${\bf A}$, and ${\bf\Phi}={\bf R}_1\cdots,{\bf R}_k$ contains the corresponding eigenvectors.

Here are the steps:

* Find the off-diagonal element $a_{ij}$ ($i\ne j$) of the greatest absolute value, called the *pivot*, and find $w=(a_{jj}-a_{ii})/2a_{ij}$.
* Solve the quadratic equation $t^2+2wt-1=0$ for $t=\tan\theta$.
* Obtain $c=1/\sqrt{1+t^2}$ and $s=t/\sqrt{1+t^2}$.
* Update all elements in the ith and jth rows and columns.

When all off-diagonal elements are zero, we get a diagonal matrix:

\begin{displaymath}{\bf\Phi}^T{\bf A}{\bf\Phi}={\bf\Lambda} \end{displaymath}

where ${\bf\Phi}={\bf R}_1{\bf R}_2{\bf R}_3\cdots$ is the product of all the rotation matrices used in the process.

**Example:** Solve the eigenvalue problem of the following matrix:

\begin{displaymath}{\bf A}=\left[\begin{array}{cc}3&2 2&1\end{array}\right] \end{displaymath}

Find the eigenvalue matrix ${\bf\Lambda}$ and eigenvector matrix ${\bf\Phi}$ so that

\begin{displaymath}{\bf\Phi}^T{\bf A}{\bf\Phi}={\bf\Lambda} \end{displaymath}

* Find $w$:

\begin{displaymath}w=\frac{a_{22}-a_{11}}{2a_{12}}=\frac{1-3}{4}=-\frac{1}{2} \end{displaymath}

* Find $t=\tan\theta$. As $w<0$, we have

\begin{displaymath}t=\tan\theta=-(w+\sqrt{w^2+1})=\frac{1}{2}(1-\sqrt{5})=-0.618 \end{displaymath}

The rotation angle is $\theta=\tan^{-1}(-0.618)=-31.7175^\circ$.

* Find $c$ and $s$

\begin{displaymath}c=\frac{1}{\sqrt{1+t^2}}=0.8507,\;\;\;\;\;\;s=\frac{t}{\sqrt{1+t^2}}=-0.5257 \end{displaymath}

* Find $\lambda_1$ and $\lambda_2$

\begin{displaymath}\lambda_1=a'_{11}=a_{11}-t a_{12}=4.2361,\;\;\;\;\;\;
\lambda_2=a'_{22}=a_{11}+t a_{12}=-0.2361 \end{displaymath}

* Find ${\bf\Phi}$ and ${\bf\Lambda}$

\begin{displaymath}
{\bf\Phi}=\left[\begin{array}{rr}c&s -s&c\end{array}\righ...
...\left[\begin{array}{cc}4.2361&0 0&-0.2361\end{array}\right]
\end{displaymath}