Eigenvalues and Eigenvectors of Matrices

Hsiu-Hau Lin
hsiuhau@phys.nthu.edu.tw
(Apr 8, 2010)

The notes cover the method to obtain eigenvalues and eigenvectors for matrices (Boas 3.11).

• Eigenvalues and eigenvectors

A matrix maps one vector to another. However, it is possible that, when acting on specific vectors, the matrix returns the same vector with a scaling factor in front,

$$M\boldsymbol{r} = \lambda \boldsymbol{r}. \tag{1}$$

The vector $\boldsymbol{r}$ is called the *eigenvector* of the matrix with the *eigenvalue* λ. Let's start with a simple example,

$$\begin{pmatrix} 5 & -2 \\ -2 & 2 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \lambda \begin{pmatrix} x \\ y \end{pmatrix}.$$

Obviously, a trivial solution $(x, y) = (0, 0)$ satisfies the above equation. To secure non-trivial solutions, the determinant must vanish and delivers the eigenvalues in return,

$$\begin{vmatrix} 5-\lambda & -2 \\ -2 & 2-\lambda \end{vmatrix} = 0 \quad \rightarrow \quad \lambda = 1, 6.$$

Once the eigenvalues are known, one can work out the corresponding eigenvectors one by one,

$$\lambda = 1, \quad \frac{1}{\sqrt{5}} \begin{pmatrix} 1 \\ 2 \end{pmatrix}; \qquad \lambda = 6, \quad \frac{1}{\sqrt{5}} \begin{pmatrix} -2 \\ 1 \end{pmatrix}.$$

One may notice that the eigenvectors form an orthonormal basis.

• Similarity transformation

The diagonal matrix is related to the original one by similarity transformation. Construct the orthogonal matrix S from the eigenvectors and the diagonal matrix D from the eigenvalues,

$$S = \begin{pmatrix} \boldsymbol{v}_1 & \boldsymbol{v}_2 \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{5}} & \frac{-2}{\sqrt{5}} \\ \frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \end{pmatrix}, \quad D = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 6 \end{pmatrix}. \tag{2}$$

The eigen equations are equivalent to the matrix identity,

$$S^{-1}MS = D. \tag{3}$$

Transformation of this kind is named *similarity transformation*. The transformation S correspond to the coordinate transformation,

$$\boldsymbol{r} = S\boldsymbol{r}'. \tag{4}$$

Thus, the original matrix M is transformed into a much simpler form in the new coordinate system,

$$M\boldsymbol{r} = \boldsymbol{R}, \quad \rightarrow \quad \left(S^{-1}MS\right)\boldsymbol{r}' = \boldsymbol{R}'.$$

That's why we go through the troubles to diagonalize the matrix and find its eigenvalues and eigenvectors.

Bringing the matrix into diagonal form can be helpful in many aspects. For instance, consider the 3×3 matrix,

$$M = \begin{pmatrix} 1 & -4 & 2 \\ -4 & 1 & -2 \\ 2 & -2 & -2 \end{pmatrix},$$

with eigenvalues $\lambda = 6, -3, -3$ and eigenvectors $(2, -2, 1)$, $(1, 1, 0)$ and $(-1, 1, 4)$. The product of M to arbitrary power can be evaluated by transforming into the eigenbasis,

$$M^n = S^{-1}D^nS.$$

It is trivial to evaluate D^n first. Then, M^n can be found by similarity transformation back to the original coordinate system. The eigenbasis also help to establish useful identity such as

$$\det\left(e^M\right) = e^{\mathrm{Tr}M}. \tag{5}$$

In the eigenbasis, the identity can be easily proved, $\det(e^M) = e^{\lambda_1}e^{\lambda_2}...e^{\lambda_n} = e^{\sum_i \lambda_i} = \exp(\mathrm{Tr}M)$. We can then use similarity transformation to arbitrary coordinate systems and establish the generality of the identity.

• Diagonalizing Hermitian matrices

Matrices are complex in general. Hermitian matrices, $H = H^\dagger$, are commonly found in physics with many surprisingly beautiful properties. For instance, the angular momentum along the z-axis is captured by a Hermitian matrix,

$$L_z = \begin{pmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

Following the same method, the Hermitian matrix can be brought into diagonal form with eigenvalues and eigenvectors,

$$\lambda = 1, 0, -1; \quad \boldsymbol{r}_1 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ i \\ 0 \end{pmatrix}, \quad \boldsymbol{r}_0 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \quad \boldsymbol{r}_{-1} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -i \\ 0 \end{pmatrix}.$$

The eigenvalues are *real* even though the matrix L_z is not. The eigenvectors with different eigenvalues are orthogonal to each other. These nice properties are true for general Hermitian matrices.

Let me walk you through the proof in the following. Write down the eigen equation and its adjoint (transpose and complex conjugate) for the Hermitian matrix,

$$H\boldsymbol{r} = \lambda \boldsymbol{r}, \qquad \boldsymbol{r}^\dagger H = \lambda^* \boldsymbol{r}^\dagger.$$

Compute inner product for the two equations, it leads to

$$\boldsymbol{r}^\dagger H \boldsymbol{r} = \lambda \boldsymbol{r}^\dagger \boldsymbol{r} = \lambda^* \boldsymbol{r}^\dagger \boldsymbol{r} \quad \to \quad (\lambda - \lambda^*) \boldsymbol{r}^\dagger \boldsymbol{r} = 0.$$

Since $\boldsymbol{r}^\dagger \boldsymbol{r} \neq 0$, the eigenvalue must be real, $\lambda = \lambda^*$. Similarly, we can play the same trick to eigen equations with different eigenvalues,

$$\boldsymbol{r}_1^\dagger H = \lambda_1 \boldsymbol{r}_1^\dagger, \qquad H\boldsymbol{r}_2 = \lambda_2 \boldsymbol{r}_2.$$

Taking the inner product, we obtain an important identity

$$\boldsymbol{r}_1^\dagger H \boldsymbol{r}_2 = \lambda_1 \boldsymbol{r}_1^\dagger \boldsymbol{r}_2 = \lambda_2 \boldsymbol{r}_1^\dagger \boldsymbol{r}_2, \quad \to \quad (\lambda_1 - \lambda_2) \boldsymbol{r}_1^\dagger \boldsymbol{r}_2 = 0.$$

Thus, if $\lambda_1 \neq \lambda_2$, the corresponding eigenvectors are orthogonal, $\boldsymbol{r}_1^\dagger \boldsymbol{r}_2 = 0$. *An interesting subtlety is left for your exercise: what happens if $\lambda_1 = \lambda_2$?*

• Lorentz transformation revisited

We now revisit the Lorentz transformation in the reduced dimensions (t, x),

$$\begin{pmatrix} ct' \\ x' \end{pmatrix} = \begin{pmatrix} \cosh\alpha & -\sinh\alpha \\ -\sinh\alpha & \cosh\alpha \end{pmatrix} \begin{pmatrix} ct \\ x \end{pmatrix}. \tag{6}$$

One can diagonalize the Lorentz transformation and obtain the eigenvalues and eigenvectors,

$$\lambda_\pm = \cosh\alpha \pm \sinh\alpha = \sqrt{\frac{c+u}{c-u}}, \sqrt{\frac{c-u}{c+u}}, \qquad \boldsymbol{r}_\pm = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ \mp 1 \end{pmatrix}. \tag{7}$$

The physical meaning for the eigenvectors are transparent: these represent the light traveling along the positive and negative x-axis. What is the physical meaning for these eigenvalues? Let me state without proof that the angular frequency ω and the wave number k follows the same Lorentz transformation,

$$\begin{pmatrix} \omega' \\ ck' \end{pmatrix} = \begin{pmatrix} \cosh\alpha & -\sinh\alpha \\ -\sinh\alpha & \cosh\alpha \end{pmatrix}\begin{pmatrix} \omega \\ ck \end{pmatrix}.$$

For light waves, $\omega = ck$ and the transformation becomes rather simple,

$$\begin{pmatrix} \omega' \\ \omega' \end{pmatrix} = \begin{pmatrix} \cosh\alpha & -\sinh\alpha \\ -\sinh\alpha & \cosh\alpha \end{pmatrix}\begin{pmatrix} \omega \\ \omega \end{pmatrix} = \sqrt{\frac{c-u}{c+u}}\begin{pmatrix} \omega \\ \omega \end{pmatrix}.$$

The relation between the angular frequencies in the moving frame and the rest frame is

$$\omega' = \sqrt{\frac{c-u}{c+u}}\omega. \tag{8}$$

The meaning of the eigenvalues is just the relativistic Doppler effect!

• Simultaneous diagonalization

Finally, I would like to comment on the eigensystems for two commuting matrices F and G. Suppose the eigen equation for the matrix F is

$$F\boldsymbol{r} = \lambda_f \boldsymbol{r}. \tag{9}$$

Acting the commuting matrix G on the equation,

$$GF\boldsymbol{r} = \lambda_f G\boldsymbol{r} \quad \rightarrow \quad F(G\boldsymbol{r}) = \lambda_f (G\boldsymbol{r}),$$

and one finds that $G\boldsymbol{r}$ is also an eigenvector with the same eigenvalue λ_f. If the eigenvalue is not degenerate, it directly implies that the eigenvectors $\boldsymbol{r}$ and $G\boldsymbol{r}$ are linearly dependent, i.e.

$$G\boldsymbol{r} = \lambda_g \boldsymbol{r}. \tag{10}$$

The above eigen equation means that F and G share the same eigenvector $\boldsymbol{r}$, but with different eigenvalues λ_f and λ_g.