Peeter Joot
peeterjoot@protonmail.com

Variational principle with two by two symmetric matrix

I pulled [1], one of too many lonely Dover books, off my shelf and started reading the review chapter. It posed the following question, which I thought had an interesting subquestion.

Exercise 1.1 Variational principle with two by two symmetric matrix.

Consider a 2×2 real symmetric matrix operator $\mathbf{O}$, with an arbitrary normalized trial vector

$$\mathbf{c} = \begin{bmatrix} \cos\theta \\ \sin\theta \end{bmatrix}. \tag{1.1}$$

The variational principle requires that minimum value of $\omega(\theta) = \mathbf{c}^\dagger \mathbf{O} \mathbf{c}$ is greater than or equal to the lowest eigenvalue.

1. If that minimum value occurs at $\omega(\theta_0)$, show that this is exactly equal to the lowest eigenvalue.

2. Explain why this is should have been anticipated.

Answer for Exercise 1.1

Part 1. If the operator representation is

$$\mathbf{O} = \begin{bmatrix} a & b \\ b & d \end{bmatrix}, \tag{1.2}$$

then the variational product is

$$\begin{aligned}
\omega(\theta) &= \begin{bmatrix} \cos\theta & \sin\theta \end{bmatrix} \begin{bmatrix} a & b \\ b & d \end{bmatrix} \begin{bmatrix} \cos\theta \\ \sin\theta \end{bmatrix} \\
&= \begin{bmatrix} \cos\theta & \sin\theta \end{bmatrix} \begin{bmatrix} a\cos\theta + b\sin\theta \\ b\cos\theta + d\sin\theta \end{bmatrix} \\
&= a\cos^2\theta + 2b\sin\theta\cos\theta + d\sin^2\theta \\
&= a\cos^2\theta + b\sin(2\theta) + d\sin^2\theta.
\end{aligned} \tag{1.3}$$

The minimum is given by

$$
\begin{aligned}
0 &= \frac{d\omega}{d\theta} \\
&= -2a \sin\theta \cos\theta + 2b \cos(2\theta) + 2d \sin\theta \cos\theta \\
&= 2b \cos(2\theta) + (d - a) \sin(2\theta),
\end{aligned}
\tag{1.4}
$$

so the extreme values will be found at

$$
\tan(2\theta_0) = \frac{2b}{a - d}. \tag{1.5}
$$

Solving for $\cos(2\theta_0)$, with $\alpha = 2b/(a - d)$, we have

$$
1 - \cos^2(2\theta) = \alpha^2 \cos^2(2\theta), \tag{1.6}
$$

or

$$
\begin{aligned}
\cos^2(2\theta_0) &= \frac{1}{1 + \alpha^2} \\
&= \frac{1}{1 + 4b^2/(a - d)^2} \\
&= \frac{(a - d)^2}{(a - d)^2 + 4b^2}.
\end{aligned}
\tag{1.7}
$$

So,

$$
\begin{aligned}
\cos(2\theta_0) &= \frac{\pm(a - d)}{\sqrt{(a - d)^2 + 4b^2}} \\
\sin(2\theta_0) &= \frac{\pm 2b}{\sqrt{(a - d)^2 + 4b^2}},
\end{aligned}
\tag{1.8}
$$

Substituting this back into $\omega(\theta_0)$ is a bit tedious. I did it once on paper, then confirmed with Mathematica (quantumchemistry/twoByTwoSymmetricVariation.nb). The end result is

$$
\omega(\theta_0) = \frac{1}{2}\left(a + d \pm \sqrt{(a - d)^2 + 4b^2}\right). \tag{1.9}
$$

The eigenvalues of the operator are given by

$$
\begin{aligned}
0 &= (a - \lambda)(d - \lambda) - b^2 \\
&= \lambda^2 - (a + d)\lambda + ad - b^2 \\
&= \left(\lambda - \frac{a + d}{2}\right)^2 - \left(\frac{a + d}{2}\right)^2 + ad - b^2 \\
&= \left(\lambda - \frac{a + d}{2}\right)^2 - \frac{1}{4}\left((a - d)^2 + 4b^2\right),
\end{aligned}
\tag{1.10}
$$

so the eigenvalues are exactly the values eq. (1.9) as stated by the problem statement.

Part 2. If the eigenvectors are $\mathbf{e}_1, \mathbf{e}_2$, the operator can be diagonalized as

$$\mathbf{O} = UDU^{\mathrm{T}}, \tag{1.11}$$

where $U = \begin{bmatrix} \mathbf{e}_1 & \mathbf{e}_2 \end{bmatrix}$, and D has the eigenvalues along the diagonal. The energy function ω can now be written

$$\begin{aligned} \omega &= \mathbf{c}^{\mathrm{T}} UDU^{\mathrm{T}} \mathbf{c} \\ &= (U^{\mathrm{T}}\mathbf{c})^{\mathrm{T}} DU^{\mathrm{T}}\mathbf{c}. \end{aligned} \tag{1.12}$$

We can show that the transformed vector $U^{\mathrm{T}}\mathbf{c}$ is still a unit vector

$$\begin{aligned} U^{\mathrm{T}}\mathbf{c} &= \begin{bmatrix} \mathbf{e}_1^{\mathrm{T}} \\ \mathbf{e}_2^{\mathrm{T}} \end{bmatrix} \mathbf{c} \\ &= \begin{bmatrix} \mathbf{e}_1^{\mathrm{T}}\mathbf{c} \\ \mathbf{e}_2^{\mathrm{T}}\mathbf{c} \end{bmatrix}, \end{aligned} \tag{1.13}$$

so

$$\begin{aligned} \left| U^{\mathrm{T}}\mathbf{c} \right|^2 &= \mathbf{c}^{\mathrm{T}}\mathbf{e}_1\mathbf{e}_1^{\mathrm{T}}\mathbf{c} + \mathbf{c}^{\mathrm{T}}\mathbf{e}_2\mathbf{e}_2^{\mathrm{T}}\mathbf{c} \\ &= \mathbf{c}^{\mathrm{T}} \left(\mathbf{e}_1\mathbf{e}_1^{\mathrm{T}} + \mathbf{e}_2\mathbf{e}_2^{\mathrm{T}} \right) \mathbf{c} \\ &= \mathbf{c}^{\mathrm{T}}\mathbf{c} \\ &= 1, \end{aligned} \tag{1.14}$$

so the transformed vector can be written as

$$U^{\mathrm{T}}\mathbf{c} = \begin{bmatrix} \cos\phi \\ \sin\phi \end{bmatrix}, \tag{1.15}$$

for some ϕ. With such a representation we have

$$\begin{aligned} \omega &= \begin{bmatrix} \cos\phi & \sin\phi \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} \begin{bmatrix} \cos\phi \\ \sin\phi \end{bmatrix} \\ &= \begin{bmatrix} \cos\phi & \sin\phi \end{bmatrix} \begin{bmatrix} \lambda_1\cos\phi \\ \lambda_2\sin\phi \end{bmatrix} \\ &= \lambda_1\cos^2\phi + \lambda_2\sin^2\phi. \end{aligned} \tag{1.16}$$

This has it's minimums where $0 = \sin(2\phi)(\lambda_2 - \lambda_1)$. For the non-degenerate case, two zeros at $\phi = n\pi/2$ for integral n. For $\phi = 0, \pi/2$, we have

$$\mathbf{c} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix}. \tag{1.17}$$

We see that the extreme values of ω occur when the trial vectors $\mathbf{c}$ are eigenvectors of the operator.

Bibliography

[1] Attila Szabo and Neil S Ostlund. *Modern quantum chemistry: introduction to advanced electronic structure theory*. Dover publications, 1989. 1