Generalized Eigenvalue Problem

The generalized eigenvalue problem of two symmetric matrices $\mathbf{A} = \mathbf{A}^T$ and $\mathbf{B} = \mathbf{B}^T$ is to find a scalar λ and the corresponding vector Φ for the following equation to hold:

$$\mathbf{A}\Phi_i = \lambda_i \mathbf{B}\Phi_i, \qquad (i = 1, \cdots, n) \tag{1}$$

or in matrix form

$$\mathbf{A}\Phi = \mathbf{B}\Phi\Lambda \tag{2}$$

where the matrixies represent

- Λ = eigenvalue
- Φ = eigenvector
- **A** = **K** matirx in the final problem
- **B** = **M** matirx in the final problem

The eigenvalue and eigenvector matrices Λ and Φ can be found in the following steps:

1 Solve the eigenvalue problem of B

Solve the eigenvalue problem of **B** to find its diagonal eigenvalue matrix Λ_B and orthogonal eigenvector matrix

$$\Phi_B = (\Phi_B^T)^{-1} \tag{3}$$

so that

$$\mathbf{B}\Phi_B = \Phi_B \Lambda_B \tag{4}$$

or

$$\Phi_B^{-1}\mathbf{B}\Phi_B = \Phi_B^T \mathbf{B}\Phi_B = \Lambda_B \tag{5}$$

2 Multiplying both sides of the second equation above by $\Lambda^{-1/2}$

Left and right multiplying both sides of the second equation above by $\Lambda^{-1/2}$ (whitening) we get

$$\Lambda_B^{-1/2}(\Phi_B^T \mathbf{B}\Phi_B)\Lambda_B^{-1/2} = \Lambda_B^{-1/2}\Lambda_B\Lambda_B^{-1/2} = \mathbf{I} \tag{6}$$

We define

$$\Phi'_B = \Phi_B \Lambda_B^{-1/2} \tag{7}$$

and get

$$(\Phi'_B)^T \mathbf{B}\Phi'_B = \mathbf{I} \tag{8}$$

Note that Φ'_B is not orthogonal

$$(\Phi')_B^{-1} = (\Phi_B \Lambda_B^{-1/2})^{-1} = \Lambda_B^{1/2}\Phi_B^{-1} = \Lambda_B^{1/2}\Phi_B^T \neq \Lambda_B^{-1/2}\Phi_B^T = \Phi_B^T \tag{9}$$

3 Apply the same transform to A

Apply the same transform to $\mathbf{A}$:

$$(\Phi_B')^T \mathbf{A} \Phi_B' = (\Lambda_B^{-1/2} \Phi_B^T) \mathbf{A} (\Phi_B \Lambda_B^{-1/2}) = \mathbf{A}' \tag{10}$$

Note that $\mathbf{A}'$ is symmetric as well as $\mathbf{A}$:

$$\mathbf{A}'^T = (\Phi_B'^T \mathbf{A} \Phi_B')^T = \Phi_B'^T \mathbf{A} \Phi_B' = \mathbf{A}' \tag{11}$$

4 Diagonalize $\mathbf{A}'$

As $\mathbf{A}'$ is symmetric,, it can be diagonalized by its orthogonal eigenvector matrix Φ_A:

$$\Phi_A^T \mathbf{A}' \Phi_A = \Lambda \tag{12}$$

i.e.,

$$\Phi_A^T (\Lambda_B^{-1/2} \Phi_B^T \mathbf{A} \Phi_B \Lambda_B^{-1/2}) \Phi_A = (\Phi_A^T \Lambda_B^{-1/2} \Phi_B^T) \mathbf{A} (\Phi_B \Lambda_B^{-1/2} \Phi_A) = \Phi^T \mathbf{A} \Phi = \Lambda \tag{13}$$

where we have defined

$$\Phi = \Phi_B \Lambda_B^{-1/2} \Phi_A \tag{14}$$

which is not orthogonal:

$$\Phi^{-1} = (\Phi_B \Lambda_B^{-1/2} \Phi_A)^{-1} = \Phi_A^T \Lambda_B^{-1/2} \Phi_B^{-1} = \Phi_A^T \Lambda_B^{1/2} \Phi_B^T \neq \Phi_A^T \Lambda_B^{-1/2} \Phi_B^T = \Phi^T \tag{15}$$

5 This Φ also diagonalizes B:

This Φ also diagonalizes $\mathbf{B}$:

$$\Phi^T \mathbf{B} \Phi = (\Phi_B \Lambda_B^{-1/2} \Phi_A)^T \mathbf{B} (\Phi_B \Lambda_B^{-1/2} \Phi_A) = \Phi_A^T (\Phi_B^T \Lambda_B^{-1/2} \mathbf{B} \Phi_B) \Lambda_B^{-1/2} \Phi_A = \Phi_A^T \Lambda_B^{-1/2} \Lambda_B \Lambda_B^{-1/2} \Phi_A = \Phi_A^T \Phi_A = \mathbf{I} \tag{16}$$

6 Now we have

Now we have

$$\begin{cases} \Phi^T \mathbf{A} \Phi = \Lambda \\ \Phi^T \mathbf{B} \Phi = \mathbf{I} \end{cases} \tag{17}$$

Right multiplying both sides of the second equation by Λ and equating the left-hand side to that of the first equation, we get

$$\mathbf{A} \Phi = \mathbf{B} \Phi \Lambda \tag{18}$$

i.e., Λ and Φ are the eigenvalue and eigenvector matrices of the generalized eigenvalue problem. Note, however, as shown above, Φ is not orthogonal. The Rayleigh quotient of two symmetric matrices $\mathbf{A}$ and $\mathbf{B}$ is a function of a vector $\mathbf{w}$ defined as:

$$R(\mathbf{w}) = \frac{\mathbf{w}^T \mathbf{A} \mathbf{w}}{\mathbf{w}^T \mathbf{B} \mathbf{w}} \tag{19}$$

To find the optimal $\mathbf{w}$ corresponding to the extremum (maximum or minimum) of $R(\mathbf{w})$, we find its derivative with respect to $\mathbf{w}$:

$$\frac{d}{d\mathbf{w}} R(\mathbf{w}) = \frac{2\mathbf{A}\mathbf{w}(\mathbf{w}^T \mathbf{B} \mathbf{w}) - 2\mathbf{B}\mathbf{w}(\mathbf{w}^T \mathbf{A} \mathbf{w})}{(\mathbf{w}^T \mathbf{B} \mathbf{w})^2} \tag{20}$$

Setting it to zero we get

$$\mathbf{A}\mathbf{w}(\mathbf{w}^T\mathbf{B}\mathbf{w}) = \mathbf{B}\mathbf{w}(\mathbf{w}^T\mathbf{A}\mathbf{w}) \tag{21}$$

i.e.,

$$\mathbf{A}\mathbf{w}\frac{\mathbf{w}^T\mathbf{A}\mathbf{w}}{\mathbf{w}^T\mathbf{B}\mathbf{w}}\mathbf{B}\mathbf{w} = R(\mathbf{w})\mathbf{B}\mathbf{w} = \lambda\mathbf{B}\mathbf{w} \tag{22}$$

The second equation can be recognized as a generalized eigenvalue problem with $\lambda = R(\mathbf{w})$ being the eigenvalue and and $\mathbf{w}$ the corresponding eigenvector. Solving this we get the vector $\mathbf{w} = \Phi$ corresponding to the maximum/minimum eigenvalue $\lambda = R(\mathbf{w})$, which maximizes/minimizes the Rayleigh quotient. next up previous