

1 Standard: Number of Channels > Number of Filters

I need to update the Cholesky decomposition of the following matrix:

$$\rho \mathbf{I} + \mathbf{D}^H \mathbf{D} \quad (1)$$

where the low-rank update takes the form:

$$\mathbf{D}^{(k+1)} = \mathbf{D}^{(k)} + \mathbf{u} \mathbf{v}^H \quad (2)$$

The low-rank update produces the following equation.

$$\rho \mathbf{I} + (\mathbf{D} + \mathbf{u} \mathbf{v}^H)^H (\mathbf{D} + \mathbf{u} \mathbf{v}^H) = \rho \mathbf{I} + \mathbf{D}^H \mathbf{D} + \mathbf{u}^H \mathbf{u} \mathbf{v} \mathbf{v}^H + \mathbf{v} \mathbf{u}^H \mathbf{D} + \mathbf{D}^H \mathbf{u} \mathbf{v}^H \quad (3)$$

The last two terms are nonHermitian, but their sum is Hermitian. The eigendecomposition of their sum expands the sum into two Hermitian terms. So, my task is to compute the eigendecomposition¹ of this expression.

$$\begin{bmatrix} \mathbf{v} & \mathbf{D}^H \mathbf{u} \end{bmatrix} \begin{bmatrix} \mathbf{D}^H \mathbf{u} & \mathbf{v} \end{bmatrix}^H = \mathbf{v} \mathbf{u}^H \mathbf{D} + \mathbf{D}^H \mathbf{u} \mathbf{v}^H \quad (4)$$

While matrix products are not commutative, some of the eigenvalues of matrix products are commutative.

Furthermore, the eigenvectors of $\mathbf{A} \mathbf{B}$ and $\mathbf{B} \mathbf{A}$ are related:

$$\mathbf{B} \mathbf{A} \mathbf{x} = \lambda \mathbf{x} \implies \mathbf{A} \mathbf{B} \mathbf{A} \mathbf{x} = \lambda \mathbf{A} \mathbf{x} \quad (5)$$

So, if \mathbf{x} is an eigenvector of $\mathbf{B} \mathbf{A}$, $\mathbf{A} \mathbf{x}$ is an eigenvector of $\mathbf{A} \mathbf{B}$.

$$\begin{bmatrix} \mathbf{D}^H \mathbf{u} & \mathbf{v} \end{bmatrix}^H \begin{bmatrix} \mathbf{v} & \mathbf{D}^H \mathbf{u} \end{bmatrix} = \begin{bmatrix} \mathbf{u}^H \mathbf{D} \mathbf{v} & \mathbf{u}^H \mathbf{D} \mathbf{D}^H \mathbf{u} \\ \mathbf{v}^H \mathbf{v} & \mathbf{v}^H \mathbf{D}^H \mathbf{u} \end{bmatrix} \quad (6)$$

The eigenvalues and corresponding eigenvectors of a 2×2 matrix can be computed using the quadratic formula. Assuming that the 2×2 matrix has 2 distinct eigenvalues, the expressions for these are below.²

$$\text{eigvals}\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right) = \frac{a+d}{2} \pm \sqrt{\left(\frac{a-d}{2}\right)^2 + bc} \quad (7)$$

$$\text{eigvals}\left(\begin{bmatrix} a & b \\ c & a^* \end{bmatrix}\right) = \text{real}(a) \pm \sqrt{bc - (\text{imag}(a))^2} \quad (8)$$

$$\text{eigvecs}\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right) = \begin{bmatrix} b \\ \frac{d-a}{2} \pm \sqrt{\left(\frac{a-d}{2}\right)^2 + bc} \end{bmatrix} \quad (9)$$

¹More specifically, I need to compute an eigendecomposition with orthogonal eigenvectors, so that the terms of the decomposition expansion are Hermitian. The matrix is Hermitian, so the eigendecomposition with orthogonal eigenvectors exists.

²It is not guaranteed the 2×2 matrix will have 2 distinct eigenvalues. In the practical considerations section, I consider those cases.

$$\text{eigvecs}\left(\begin{bmatrix} a & b \\ c & a^* \end{bmatrix}\right) = \begin{bmatrix} b \\ -j \text{imag}(a) \pm \sqrt{bc - (\text{imag}(a))^2} \end{bmatrix} \quad (10)$$

$$\text{eigvals}\left(\begin{bmatrix} \mathbf{u}^H \mathbf{D} \mathbf{v} & \mathbf{u}^H \mathbf{D} \mathbf{D}^H \mathbf{u} \\ \mathbf{v}^H \mathbf{v} & \mathbf{v}^H \mathbf{D}^H \mathbf{u} \end{bmatrix}\right) = \text{real}(\mathbf{u}^H \mathbf{D} \mathbf{v}) \pm \sqrt{\mathbf{v}^H \mathbf{v} \mathbf{u}^H \mathbf{D} \mathbf{D}^H \mathbf{u} - (\text{imag}(\mathbf{u}^H \mathbf{D} \mathbf{v}))^2} \quad (11)$$

$$\text{eigvecs}\left(\begin{bmatrix} \mathbf{u}^H \mathbf{D} \mathbf{v} & \mathbf{u}^H \mathbf{D} \mathbf{D}^H \mathbf{u} \\ \mathbf{v}^H \mathbf{v} & \mathbf{v}^H \mathbf{D}^H \mathbf{u} \end{bmatrix}\right) = \begin{bmatrix} \mathbf{u}^H \mathbf{D} \mathbf{D}^H \mathbf{u} \\ -j \text{imag}(\mathbf{u}^H \mathbf{D} \mathbf{v}) \pm \sqrt{\mathbf{v}^H \mathbf{v} \mathbf{u}^H \mathbf{D} \mathbf{D}^H \mathbf{u} - (\text{imag}(\mathbf{u}^H \mathbf{D} \mathbf{v}))^2} \end{bmatrix} \quad (12)$$

Therefore,

$$\text{eigvecs}(\mathbf{v} \mathbf{u}^H \mathbf{D} + \mathbf{D}^H \mathbf{u} \mathbf{v}^H) = \mathbf{u}^H \mathbf{D} \mathbf{D}^H \mathbf{u} \mathbf{v} + (-j \text{imag}(\mathbf{u}^H \mathbf{D} \mathbf{v}) \pm \sqrt{\mathbf{v}^H \mathbf{v} \mathbf{u}^H \mathbf{D} \mathbf{D}^H \mathbf{u} - (\text{imag}(\mathbf{u}^H \mathbf{D} \mathbf{v}))^2}) \mathbf{D}^H \mathbf{u} \quad (13)$$

$$\text{eigvals}(\mathbf{v} \mathbf{u}^H \mathbf{D} + \mathbf{D}^H \mathbf{u} \mathbf{v}^H) = \text{real}(\mathbf{u}^H \mathbf{D} \mathbf{v}) \pm \sqrt{\mathbf{v}^H \mathbf{v} \mathbf{u}^H \mathbf{D} \mathbf{D}^H \mathbf{u} - (\text{imag}(\mathbf{u}^H \mathbf{D} \mathbf{v}))^2} \quad (14)$$

2 Transposed Version: Number of Channels < Number of Filters

Alternatively, the form

$$\rho \mathbf{I} + \mathbf{D} \mathbf{D}^H \quad (15)$$

can be used exploiting by exploiting the Woodbury Matrix Lemma. As before, a rank-1 dictionary update takes the form:

$$\mathbf{D}^{(k+1)} = \mathbf{D}^{(k)} + \mathbf{u} \mathbf{v}^H \quad (16)$$

The low-rank update produces the following equation:

$$\rho \mathbf{I} + (\mathbf{D} + \mathbf{u} \mathbf{v}^H)(\mathbf{D} + \mathbf{u} \mathbf{v}^H)^H = \rho \mathbf{I} + \mathbf{D} \mathbf{D}^H + \mathbf{v}^H \mathbf{v} \mathbf{u} \mathbf{u}^H + \mathbf{u} \mathbf{v}^H \mathbf{D}^H + \mathbf{D} \mathbf{v} \mathbf{u}^H \quad (17)$$

As in the previous case, the last two terms are nonHermitian, but their sum is Hermitian. So, my task, as before, is to compute the eigendecomposition of this expression.

$$\begin{bmatrix} \mathbf{u} & \mathbf{D} \mathbf{v} \end{bmatrix} \begin{bmatrix} \mathbf{D} \mathbf{v} & \mathbf{u} \end{bmatrix}^H = \mathbf{u} \mathbf{v}^H \mathbf{D}^H + \mathbf{D} \mathbf{v} \mathbf{u}^H \quad (18)$$

If \mathbf{x} is an eigenvector of $\mathbf{B} \mathbf{A}$, $\mathbf{A} \mathbf{x}$ is an eigenvector of $\mathbf{A} \mathbf{B}$.

$$\begin{bmatrix} D\mathbf{v} & \mathbf{u} \end{bmatrix}^H \begin{bmatrix} \mathbf{u} & D\mathbf{v} \end{bmatrix} = \begin{bmatrix} \mathbf{v}^H D^H \mathbf{u} & \mathbf{v}^H D^H D\mathbf{v} \\ \mathbf{u}^H \mathbf{u} & \mathbf{u}^H D\mathbf{v} \end{bmatrix} \quad (19)$$

The eigenvalues and corresponding eigenvectors of a 2×2 matrix can be computed using the quadratic formula.

$$\text{eigvals}\left(\begin{bmatrix} a & b \\ c & a^* \end{bmatrix}\right) = \text{real}(a) \pm \sqrt{bc - (\text{imag}(a))^2} \quad (20)$$

$$\text{eigvecs}\left(\begin{bmatrix} a & b \\ c & a^* \end{bmatrix}\right) = \begin{bmatrix} b \\ -j \text{imag}(a) \pm \sqrt{bc - (\text{imag}(a))^2} \end{bmatrix} \quad (21)$$

$$\text{eigvals}\left(\begin{bmatrix} \mathbf{v}^H D\mathbf{u} & \mathbf{v}^H D^H D\mathbf{v} \\ \mathbf{u}^H \mathbf{u} & \mathbf{u}^H D\mathbf{v} \end{bmatrix}\right) = \text{real}(\mathbf{v}^H D^H \mathbf{u}) \pm \sqrt{\mathbf{u}^H \mathbf{u} \mathbf{v}^H D^H D\mathbf{v} - (\text{imag}(\mathbf{v}^H D^H \mathbf{u}))^2} \quad (22)$$

$$\text{eigvecs}\left(\begin{bmatrix} \mathbf{v}^H D^H \mathbf{u} & \mathbf{v}^H D^H D\mathbf{v} \\ \mathbf{u}^H \mathbf{u} & \mathbf{u}^H D\mathbf{v} \end{bmatrix}\right) = \begin{bmatrix} \mathbf{v}^H D^H D\mathbf{v} \\ -j \text{imag}(\mathbf{v}^H D^H \mathbf{u}) \pm \sqrt{\mathbf{u}^H \mathbf{u} \mathbf{v}^H D^H D\mathbf{v} - (\text{imag}(\mathbf{v}^H D^H \mathbf{u}))^2} \end{bmatrix} \quad (23)$$

Therefore,

$$\text{eigvecs}(\mathbf{u} \mathbf{v}^H D^H + D\mathbf{v} \mathbf{u}^H) = \mathbf{v}^H D^H D\mathbf{v} \mathbf{u} + (-j \text{imag}(\mathbf{v}^H D^H \mathbf{u}) \pm \sqrt{\mathbf{u}^H \mathbf{u} \mathbf{v}^H D^H D\mathbf{v} - (\text{imag}(\mathbf{v}^H D^H \mathbf{u}))^2}) D\mathbf{v} \quad (24)$$

$$\text{eigvals}(\mathbf{u} \mathbf{v}^H D^H + D\mathbf{v} \mathbf{u}^H) = \text{real}(\mathbf{v}^H D^H \mathbf{u}) \pm \sqrt{\mathbf{u}^H \mathbf{u} \mathbf{v}^H D^H D\mathbf{v} - (\text{imag}(\mathbf{v}^H D^H \mathbf{u}))^2} \quad (25)$$

3 Practical Considerations

3.1 Fewer Eigenvectors

Starting with linear algebra principles:

$$\mathbf{B}\mathbf{A}\mathbf{x} = \lambda\mathbf{x} \implies \mathbf{A}\mathbf{B}(\mathbf{A}\mathbf{x}) = \lambda(\mathbf{A}\mathbf{x}) \quad (26)$$

Furthermore,

$$\mathbf{A}\mathbf{B}\mathbf{y} = \lambda\mathbf{y} \implies \mathbf{B}\mathbf{A}(\mathbf{B}\mathbf{y}) = \lambda(\mathbf{B}\mathbf{y}) \quad (27)$$

If eigenvector \mathbf{y} is in the nullspace of matrix \mathbf{B} , \mathbf{y} cannot be "discovered" by applying matrix \mathbf{A} to an eigenvector of $\mathbf{B}\mathbf{A}$.

3.1.1 Standard: Number of Channels > Number of Filters

$$\mathbf{B} = [\mathbf{D}^H \mathbf{u} \quad \mathbf{v}]^H \quad (28)$$

The matrix \mathbf{AB} is a Hermitian $M \times M$ matrix, so it has M real eigenvalues and M independent eigenvectors which can be chosen to be orthogonal. Therefore, \mathbf{BA} has 2 independent eigenvectors if \mathbf{B} is rank 2. What could cause matrix \mathbf{B} to be less than rank 2?

1.
$$\mathbf{D}^H \mathbf{u} = \mathbf{0} \quad (29)$$

2.
$$\mathbf{v} = \mathbf{0} \quad (30)$$

3.
$$\mathbf{D}^H \mathbf{u} = \alpha \mathbf{v} \quad (31)$$

In the first 2 cases, the matrix \mathbf{BA} has one independent eigenvector. The first case implies that only the Hermitian update is nonzero. The second case implies that the entire update is zero. (The eigenvalues are zero, so as long as the normalization of eigenvectors is handled with care, it is not necessary to check for these cases in code.)

In the third case, the diagonalization of \mathbf{AB} can be determined directly without using the 2×2 matrix \mathbf{BA} .

$$\mathbf{AB} = 2 \operatorname{real}(\alpha) \mathbf{v} \mathbf{v}^H \quad (32)$$

$$\lambda = 2 \operatorname{real}(\alpha) \|\mathbf{v}\|_2^2 \quad (33)$$

$$\mathbf{x} = \frac{\mathbf{v}}{\|\mathbf{v}\|_2} \quad (34)$$

The rest of the eigenvalues are zero. (The corresponding 2×2 matrix \mathbf{BA} shares the same nonzero eigenvalue. The eigenvector that is lost has an eigenvalue of zero, so like in the other 2 cases, it is not necessary to check for this case in code.)

3.1.2 Transposed: Number of Channels < Number of Filters

$$\mathbf{B} = [\mathbf{D} \mathbf{v} \quad \mathbf{u}]^H \quad (35)$$

The matrix \mathbf{AB} is a Hermitian $C \times C$ matrix, so it has C real eigenvalues and C independent eigenvectors which can be chosen to be orthogonal. Therefore, \mathbf{BA} has 2 independent eigenvectors if \mathbf{B} is rank 2. What could cause matrix \mathbf{B} to be less than rank 2?

1.
$$\mathbf{D} \mathbf{v} = \mathbf{0} \quad (36)$$

2.

$$\mathbf{u} = \mathbf{0} \quad (37)$$

3.

$$\mathbf{D}\mathbf{v} = \alpha\mathbf{u} \quad (38)$$

The first case implies that only the Hermitian update is nonzero. The second case implies that the entire update is zero.

In the third case, the diagonalization of \mathbf{AB} can be determined directly without using 2×2 matrix \mathbf{BA} .

$$\mathbf{AB} = 2 \operatorname{real}(\alpha) \mathbf{u} \mathbf{u}^H \quad (39)$$

$$\lambda = 2 \operatorname{real}(\alpha) \|\mathbf{u}\|_2^2 \quad (40)$$

$$\mathbf{x} = \frac{\mathbf{u}}{\|\mathbf{u}\|_2} \quad (41)$$

The rest of the eigenvalues are zero. (The corresponding 2×2 matrix \mathbf{BA} shares the same nonzero eigenvalue. The eigenvector that is lost has an eigenvalue of zero, so like in the other 2 cases, it is not necessary to check for this case in code.)

3.2 Identical Eigenvalues

If the pair of eigenvalues are the same, then all nonzero vectors are eigenvectors of \mathbf{BA} . However, it is necessary for a diagonalization expansion with only Hermitian terms that the eigenvectors of \mathbf{AB} are chosen to be orthogonal, which can be found using a Gram-Schmidt process. This case is not just a theoretical concern; it is necessary to check for this in code.

3.2.1 Standard: Number of Channels > Number of Filters

Suppose these eigenvectors are chosen for the 2×2 matrix:

$$\mathbf{x}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \quad (42)$$

$$\mathbf{x}_2 = \begin{bmatrix} -\frac{\mathbf{v}^H \mathbf{D}^H \mathbf{u}}{\mathbf{v}^H \mathbf{v}} \\ 1 \end{bmatrix} \quad (43)$$

The corresponding eigenvectors \mathbf{Ax}_i of matrix \mathbf{AB} are orthogonal.

3.2.2 Transposed: Number of Channels < Number of Filters

Suppose these eigenvectors are chosen for the 2×2 matrix:

$$\mathbf{x}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \quad (44)$$

$$\mathbf{x}_2 = \begin{bmatrix} -\frac{\mathbf{u}^H \mathbf{D} \mathbf{v}}{\mathbf{u}^H \mathbf{u}} \\ 1 \end{bmatrix} \quad (45)$$

The corresponding eigenvectors $\mathbf{A}\mathbf{x}_i$ of matrix $\mathbf{A}\mathbf{B}$ are orthogonal.