Advanced Eigenvalues and Eigenvectors

Wayne Dam

Physics Cafe

September 23, 2020

Outline

Transpose of Products

- Transpose (^{T}) and the Conjugate or Hermitian Transpose (†) will reverse products of matrices.
- Transposes and conjugate transposes:

$$(AB)^\dagger = B^\dagger A^\dagger \tag{1}$$

Proof.

For the transpose expand the lhs and rhs explicitly using the transposed element rule $[C^T]_{ij} = [C]_{ji}$ for a matrix C and the matrix product rule $[AB]_{ij} = \sum_k a_{ik} b_{kj}$. □

- Note that because matrix multiplication is associative this reversal applies to any number of matrices in the product. So for three matrices,

$$\begin{aligned}(ABC)^\dagger &= \Big((AB)C\Big)^\dagger && (2)\\ &= C^\dagger (AB)^\dagger && (3)\\ &= C^\dagger B^\dagger A^\dagger && (4)\end{aligned}$$

Inverse of Products

- The matrix inverse will reverse products as well.

$$(AB)^{-1} = B^{-1}A^{-1} \tag{5}$$

Proof.

We start with the definition for inverses that $CC^{-1} = I$. Therefore,

$(AB)(AB)^{-1} = I$	By definition of inverse.
$A^{-1}AB(AB)^{-1} = A^{-1}$	Multiply lhs and rhs by A^{-1}.
$B(AB)^{-1} = A^{-1}$	Cancels the A's on lhs.
$B^{-1}B(AB)^{-1} = B^{-1}A^{-1}$	Multiply both sides by B^{-1}.
$(AB)^{-1} = B^{-1}A^{-1}$	Cancel the B's on the lhs.

□

- Or think of it as putting on socks (B operation) and shoes (A operation) and remembering matrix multiplication goes left to right.

Hermitian and Unitary Operators

Definition (Hermitian Matrix)

A matrix is Hermitian when it equals its conjugate transpose (also called its Hermitian).

$$A = A^\dagger \tag{6}$$

Definition (Unitary Matrix)

A unitary matrix is defined by having its inverse equal to its conjugate transpose.

$$U^{-1} = U^\dagger \tag{7}$$

They have some remarkable properties that are key to quantum mechanics.

Definition of Eigenvalues/Eigenvectors

- Recall the original definition of eigenvalues and eigenvectors, presented on May 10'th by John.

Definition (Eigenvalue/Eigenvector of a Matrix)

Given a linear operator, A, an eigenvalue/eigenvalue pair will satisfy,

$$Au = \lambda u. \tag{8}$$

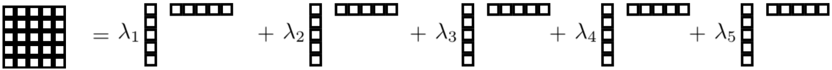
- That is, a special vector u, when acted on by A, results in a scaled version of the original vector. This is *not* true for any general vector.
- There can be more than one pair for a given A. In bra-ket notation we write,

$$A|u_i\rangle = \lambda_i|u_i\rangle. \tag{9}$$

or simply $A|i\rangle = \lambda_i|i\rangle$.

Spectral Decomposition of Hermitian Operators

- Hermitian operators have a special property in that they can be be expanded as a weighted sum of outer products of the eigenvectors. The weights are given by the eigenvalues.
- Pictorially we can sketch this decomposition as in the figure.



Theorem (Spectral Decomposition)

Any Hermitian matrix has a diagonal representation $A = \sum_i \lambda_i |i\rangle\langle i|$.

- The number of (non-zero) eigenvalues is the *rank* of the matrix.
- Writing it regular vector notation we have

$$A = \lambda_1 u_1 u_1^\dagger + \lambda_2 u_2 u_2^\dagger + \dots \tag{10}$$

$$= \sum_{i=1}^{n} \lambda_i u_i u_i^\dagger \tag{11}$$

where n is *at most* the dimension of the Hermitian matrix, A.

Properties of the Spectral Decomposition

Properties of the Spectral Decomposition

1. All eigenvalues are real.
2. All eigenvectors are orthogonal.
3. The set of eigenvectors span the whole space.

1. Proof (All eigenvalues are real).

Start with the basic definition in matrix/vector notation.

$Au = \lambda u$	Definition.	(12)
$u^\dagger A u = \lambda u^\dagger u$	Multiply sides by $u^\dagger$.	(13)
$u^\dagger A^\dagger u = \lambda^* u^\dagger u$	Take Hermitian both sides.	(14)
$u^\dagger A u = \lambda^* u^\dagger u$	Use A is Hermitian on lhs.	(15)
$\lambda u^\dagger u = \lambda^* u^\dagger u$	Use definition on lhs.	(16)
$\lambda = \lambda^*$	Cancel the scaler $u^\dagger u$.	(17)

Therefore, λ is real. $\square$

2. Proof (All eigenvectors belonging to *distinct* eigenvalues are orthogonal).

- Start with two eigenvalues $\lambda \neq \mu$ with corresponding eigenvectors u, v.
- So we have $Au = \mu u$ and $Av = \lambda v$.
- Now look at the inner product $u^\dagger v$ scaled by λ,

$$\begin{aligned}
\lambda u^\dagger v &= u^\dagger \lambda v && \text{Shift scalar } \lambda \text{ into expression.} && (18)\\
&= u^\dagger A v && \text{Use definition of } \lambda v && (19)\\
&= (A^\dagger u)^\dagger v && \text{Use the transpose of products property.} && (20)\\
&= (Au)^\dagger v && \text{Use the fact } A \text{ is Hermitian.} && (21)\\
&= (\mu u)^\dagger v && \text{Use the definition of } Au && (22)\\
&= u^\dagger \mu^* v && \text{Transpose of products again.} && (23)\\
&= \mu u^\dagger v && \text{Use the fact eigevalues are real.} && (24)
\end{aligned}$$

- Subtracting the rhs from the lhs and factoring out the $u^\dagger v$ we end up with $(\lambda - \mu) u^\dagger v = 0$.
- Since we assume λ and μ are different we have $u^\dagger v = 0$ and they are orthogonal.

□

Matrix Version of the Spectral Decomposition

- We can also write the spectral decomposition another way.
- We have $Au = \lambda u$ for each eigenvector/eigenvalue pair.
- The lhs and rhs are both vectors and we can write a matrix equation for all the eigenvalues/eigenvector pairs simultaneously by arranging the columns into a matrix on each side.

$$\begin{pmatrix} Au_1 & Au_2 & \cdots & Au_n \\ \vdots & \vdots & & \vdots \end{pmatrix} = \begin{pmatrix} \lambda_1 u_1 & \lambda_2 u_2 & \cdots & \lambda_n u_n \\ \vdots & \vdots & & \vdots \end{pmatrix} \tag{25}$$

$$A \begin{pmatrix} u_1 & u_2 & \cdots & u_n \\ \vdots & \vdots & & \vdots \end{pmatrix} = \begin{pmatrix} u_1 & u_2 & \cdots & u_n \\ \vdots & \vdots & & \vdots \end{pmatrix} \begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{pmatrix} \tag{26}$$

- This lets us write $AU = U\Lambda$.
- And so,

$$A = U\Lambda U^{-1} \tag{27}$$

Properties of U

1. The columns of U are orthonormal. Why?

2. The (i,j)'th element of the product $U^\dagger U$ is $u_i^\dagger u_j = \delta_{ij}$.
   Therefore, $U^\dagger U = I$ and $U^\dagger = U^{-1}$.

   This gives us the first property.

Property: The matrix U is Unitary

$$U^{-1} = U^\dagger \tag{28}$$

$$\implies A = U\Lambda U^\dagger \tag{29}$$

3. Because $UU^\dagger = I$ we can expand it as $\sum_i u_i u_i^\dagger = I$.
   In bra-ket notation we can write this as

Property: Resolution of the Identity

$$\sum_i |i\rangle\langle i| = I \tag{30}$$

Change of Basis for Decoupling

4. Treating A as a transformation from the vector v to y, we have,

$$\begin{aligned} y &= Av && (31)\\ &= U\Lambda U^{\dagger} v && (32)\\ &= U\Lambda \begin{pmatrix} u_1^{\dagger} \\ u_2^{\dagger} \\ \vdots \end{pmatrix} v && (33)\\ &= U\Lambda \begin{pmatrix} \mathrm{v}_1' \\ \mathrm{v}_2' \\ \vdots \end{pmatrix} = U \begin{pmatrix} \lambda_1 \mathrm{v}_1' \\ \lambda_2 \mathrm{v}_2' \\ \vdots \end{pmatrix} && (34) \end{aligned}$$

Property: Decoupling by Change of Basis

- The eigenvectors form new orthonormal coordinate system that decouples the transformation of A into independent components, each scaled by a single eigenvalue.
- U acts as a change of basis transformation. Each $u_i^{\dagger} v$ calculates the component of v projected onto the i'th new basis.

Applications of Diagonalization

Example (1. Decoupling of Variables)

- Consider the symmetric matrix $A = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$.
- It has the eigenvalues 3 and 1 and corresponding eigenvectors $\frac{1}{\sqrt{2}}(1, 1)$ and $\frac{1}{\sqrt{2}}(1, -1)$.
- The action of A as a transformation is shown below.

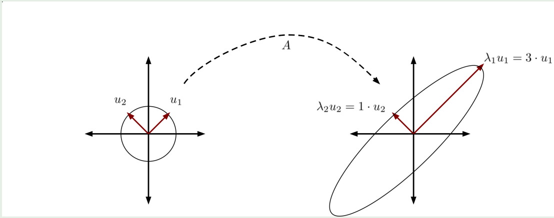


Figure: Action of A as a linear transformation. Scaling of the eigenvectors and mapping of the unit circle to an ellipse.

- The eigenvectors define the major axis of the ellipse and the eigenvalues their length.

Example (1. Decoupling of Variables, cont'd)

- Consider the quadratic form $v^\dagger A v$ where we write $v = (x, y)$.

$$\begin{pmatrix} x & y \end{pmatrix} \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} \tag{35}$$

- Expanding and simplifying this gives us the real scalar in terms of x and y.

$$2x^2 + 2xy + 2y^2 \tag{36}$$

- Note how the two variables are cross coupled.
- Imagine a problem involving hundreds of variables all cross coupled up to second order. How to make sense of it?

Example (1. Decoupling of Variables, cont'd)

- Write the quadratic form using the eigen-decomposition, $A = U\Lambda U^\dagger$.

$$\begin{pmatrix} x & y \end{pmatrix} \cdot \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \cdot \begin{pmatrix} 3 & 0 \\ 0 & 1 \end{pmatrix} \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \cdot \begin{pmatrix} x \\ y \end{pmatrix} \tag{37}$$

$$= \begin{pmatrix} \frac{x+y}{\sqrt{2}} & \frac{x-y}{\sqrt{2}} \end{pmatrix} \cdot \begin{pmatrix} 3 & 0 \\ 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} \frac{x+y}{\sqrt{2}} \\ \frac{x-y}{\sqrt{2}} \end{pmatrix} \tag{38}$$

$$= \begin{pmatrix} x' & y' \end{pmatrix} \cdot \begin{pmatrix} 3 & 0 \\ 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} x' \\ y' \end{pmatrix} \tag{39}$$

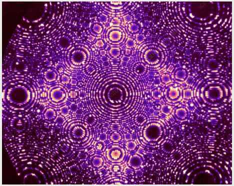
$$= 3x'^2 + y'^2 \tag{40}$$

  where we've defined the new variables $x' = \frac{x+y}{\sqrt{2}}$ and $y' = \frac{x-y}{\sqrt{2}}$.
- We've decoupled the quadratic interactions into independent components with each component scaled by an eigenvalue.
- In general with n variables the original expression involves about $n^2/2$ coupled quadratic terms. Interpretation is almost impossible.
- After the eigen-decomposition we are back to n redefined variables with no coupling, each simply scaled by a number, an eigenvalue. We can interpret the system straightforwardly in the new variables.
- E.g., *Principle Component Analysis* in statistics and machine learning.

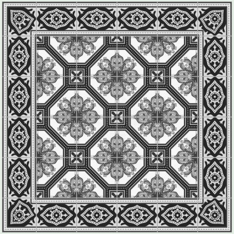
Image Compression by Low Rank Approximation

Example (2. Image Compression)

- Digital images can be considered matrices of numbers with the values representing pixel intensity.
- Consider an image that is symmetric along its diagonal so its matrix will be symmetric. For example, a crystal x-ray photo or a patterned rug.



(a) X-Ray Diffraction of Iridium.



(b) A patterned rug (800×800 pixels).

Figure: Examples of images with diagonal symmetry.

- Considered as a matrix it is perfectly valid to consider its eigen-decomposition into eigenvalues and eigenvectors in an 800-dimensional vector space.

Example (2. Image Compression, cont'd)

- Here is the original 800×800 pixel grey scale image. Each pixel has a value ranging from 0 (black) to 255 (white).

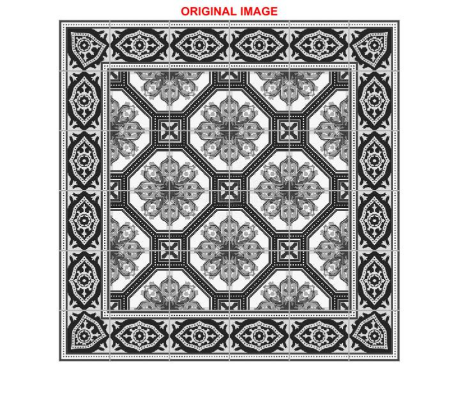


Figure: Diagonally symmetric 800×800 pixel grey scale image.

Example (2. Image Compression, cont'd)

- Treating the image as a Hermitian operator and forming the decomposition $UDU^\dagger$, we can plot the 800 real eigenvalues along the diagonal of D sorted from largest to smallest.

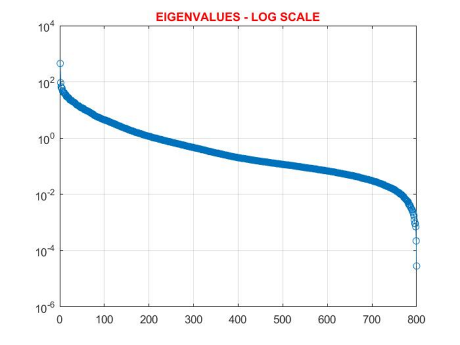
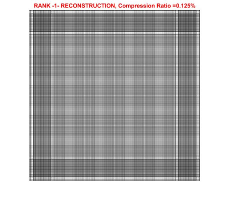


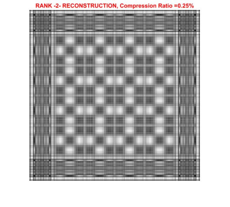
Figure: Eigenvalue spectrum of image on a log scale.

Example (2. Image Compression, cont'd)

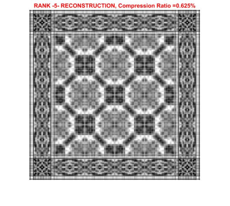
- Using the first few eigenvectors and corresponding eigenvalues of the decomposition we can perform low rank reconstructions of the original image.



(a) Rank 1



(b) Rank 2



(c) Rank 5

(d) Rank 15

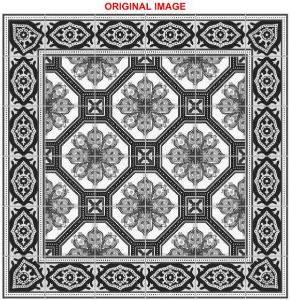
Example (2. Image Compression, cont'd)

- We can also plot the reconstruction error as a function of the reproduction rank.
- Note that the compression ratio is close to the number of columns used in the reconstruction (i.e., the rank) divided by the number of original columns (800). We can achieve very high compression ratios for a reasonable distortion.

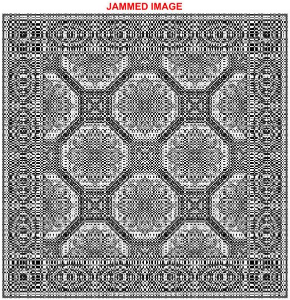
Figure: Reconstruction error versus rank.

Example (3. De-noising)

- Often a signal will have distortion due to interference or noise.



(a) Original.



(b) With interference.

Example (3. De-noising, cont'd)

- The interference often shows structure in the eigenvalue spectrum. This is analogous to interference in the frequency spectrum of a signal.

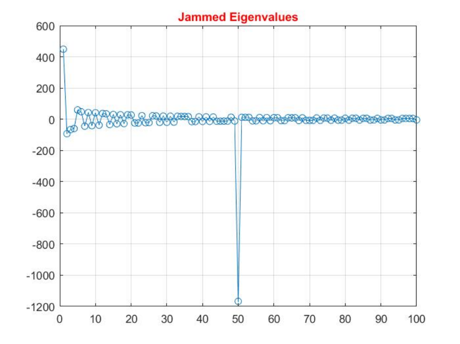
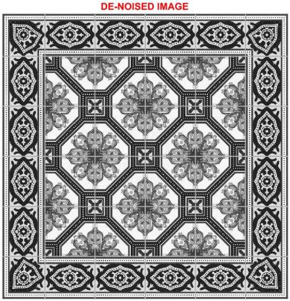


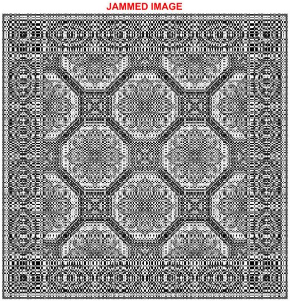
Figure: Eigenvalue spectrum of the jammed image showing distortion in the 50'th eigenvalue

Example (3. De-noising, cont'd)

- To remove the jamming we can just set eigenvalue 50 to zero. This lets us reconstruct the original image by pre and post multiplying the new eigenvalue matrix with U and $U^\dagger$.



(a) De-noised image.

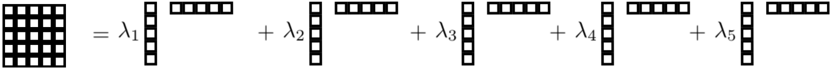


(b) With interference.

Summary

MAIN TAKE AWAYS

1. A Hermitian matrix can be decomposed into a weighted sum of rank one outer products.

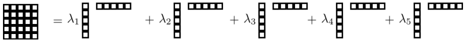


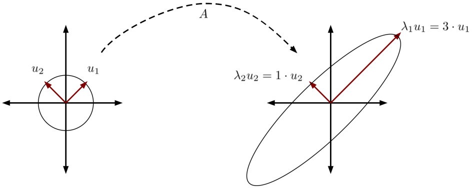
2. The eigenvectors form a complete orthonormal basis of the vector space.

3. Based on the eigenvalues we can decompose the vector space into two orthogonal linear subspaces, the *signal subspace* and the *noise subspace*.
   One space we consider relevant and one we do not.

4. This can be exploited to do useful and interesting things.

Summary, cont'd

5. There are three equivalent definitions of eigenvalues and eigenvalues:
    1. Original scaling property:
    $$Au = \lambda u. \tag{41}$$
    2. The spectral or diagonal decomposition of Hermitian matrices:
    $$A = \sum_i \lambda_i |i\rangle\langle i| \tag{42}$$
    

    $$A = U\Lambda U^\dagger \tag{43}$$
    3. The eigenvectors lie along the major and minor axis of the ellipse defined by a Hermitian matrix; the eigenvalues give the length of the axis.
    

    4. I know one more ...

End

END.