Eigen decomposition of a matrix

Mahdi Akbari Zarkesh
9612762638

What is Eigen?

Eigen means own or self. In linear algebra, eigenvalue, eigenvector and Eigen decomposition are terms that are intrinsically related. Eigen decomposition is the method to decompose a square matrix into its eigenvalues and eigenvectors. For a matrix A, if.

$$Av=\lambda v \quad (1)$$

Then v is an eigenvector of matrix A and λ is the corresponding eigenvalue. That is, if matrix A is multiplied by a vector and the result is a scaled version of the same vector, then it is an eigenvector of A and the scaling factor is its eigenvalue.

Eigendecomposition

So how do we find the eigenvectors of a matrix? From (1):

$$Av-\lambda Iv=0$$
$$(A-\lambda I)v=0 \quad (2)$$

Where I is the identity matrix. The values of λ where (2) holds are the eigenvalues of A. It turns out that this equation is equivalent to:

$$\det(A-\lambda I)=0 \quad (3)$$

Where det() is the determinant of a matrix.

Why is eigendecomposition useful?

we can join both eigenvectors and eigenvalues in a single matrix equation

$$A\left[\mathbf{v_1}\mathbf{v_2}\right] = \begin{bmatrix} v_{11} & v_{21} \\ v_{12} & v_{22} \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} = \left[\mathbf{v_1}\mathbf{v_2}\right] \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}$$

If we replace:

$$\Lambda = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}$$

$$V = \left[\mathbf{v_1}\mathbf{v_2}\right]$$

it is also true that:

$$AV = V\Lambda$$
$$A = V\Lambda V^{-1} \quad (4)$$

Eigen decomposition decomposes a matrix A into a multiplication of a matrix of eigenvectors V and a diagonal matrix of eigenvalues Λ. This can only be done if a matrix is diagonalizable. In fact, the definition of a diagonalizable matrix $A \in R_{n\times n}$ is that it can be Eigen decomposed into n eigenvectors, so that $V^{-1}AV=\Lambda$.

Power of a matrix with Eigen decomposition

From (4):

$$A^2=V\Lambda V^{-1}V\Lambda V^{-1}=V\Lambda^2 V^{-1}$$
$$A^n=V\Lambda^n V^{-1}$$

The power of Λ is just the power of each diagonal element. This becomes much simpler than multiplications of A.

Properties of Eigen decomposition

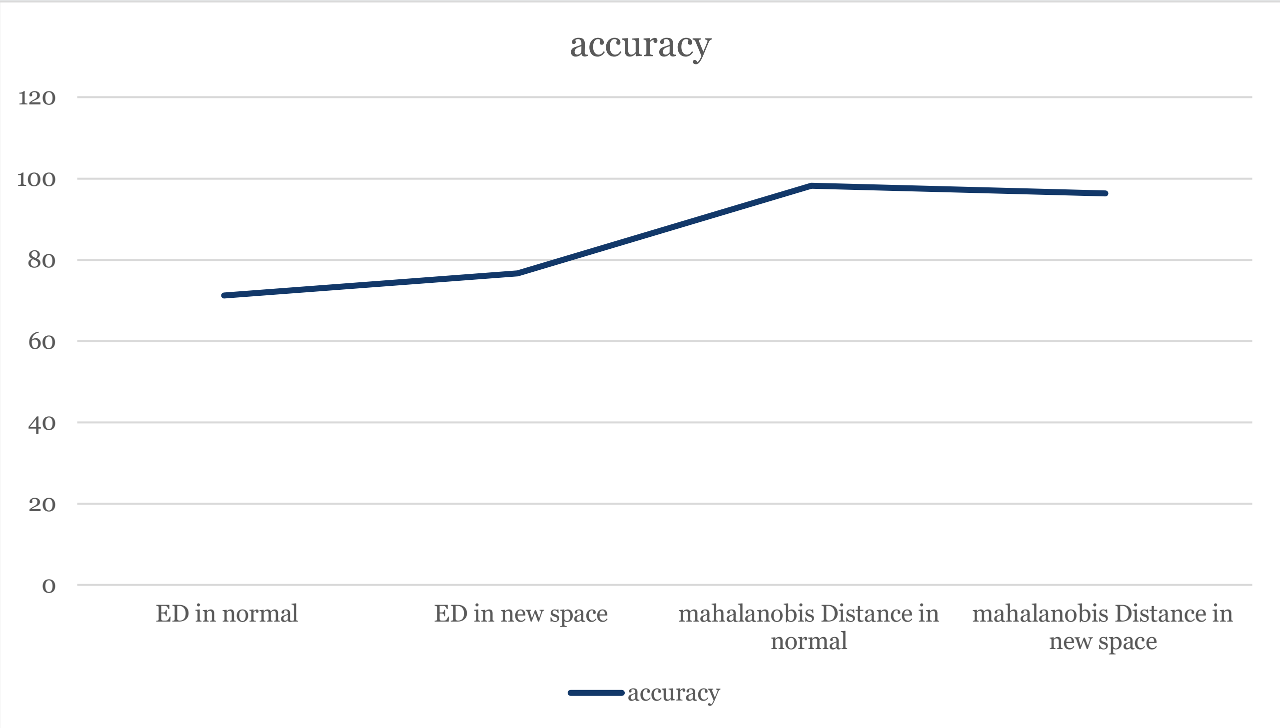
- $det(A) = \prod_{i=1}^{n} \lambda_i$ (the determinant of A is equal to the product of its eigenvalues)
- $tr(A) = \sum_{i=1}^{n} \lambda_i$ (the trace of A is equal to the sum of its eigenvalues)
- The eigenvalues of A^{-1} are λ_i^{-1}
- The eigenvalues of A^n are λ_i^n
- In general, the eigenvalues of $f(A)$ are $f(\lambda_i)$
- The eigenvectors of A^{-1} are the same as the eigenvectors of A.
- if A is hermitian (its conjugate transpose is equal to itself) and full-rank (all rows or columns are linearly independent), then the eigenvectors are mutually orthogonal (the dot-product between any two eigenvectors is zero) and the eigenvalues are real.
- A is invertible if all its eigenvalues are different from zero and vice-versa.
- If the eigenvalues of matrix A are distinct (not repeated), then A can be eigendecomposed.

More details can find https://en.wikipedia.org/wiki/QR_decomposition
And https://medium.com/machine-learning-mindset/matrix-eigendecomposition-its-importance-and-the-applications-26c92f76731

The Dataset : wine

Classes : 3
Samples per class : [59,71,48]
Samples total : 178
Dimensionality : 13
Features : real, positive

We change the space of data by Φ(x) = Q'*landa^(1/2)*x and use ED distance to see the effect of this change on our Dataset:



It's totally like using Mahalonobis distance.The Mahalanobis distance is a measure of the distance between a point P and a distribution D. It is a multi-dimensional generalization of the idea of measuring how many standard deviations away P is from the mean of D. This distance is zero if P is at the mean of D, and grows as P moves away from the mean along each principal component axis. If each of these axes is re-scaled to have unit variance, then the Mahalanobis distance corresponds to standard Euclidean distance in the transformed space. The Mahalanobis distance is thus unit less and scale-invariant, and takes into account the correlations of the data set.

Thank you :)