

Linear algebra is the study of lines and planes, vector spaces and mappings that are required for linear transforms.

MATRIX DECOMPOSITION

1. Way of reducing a matrix into its constituent parts to simplify more complex matrix operations that can be performed on the decomposed matrix rather than on the original matrix itself.

LU Decomposition

1. Used for Square Matrices(a **matrix with the same number of rows and columns**).

$$A = L \cdot U$$

A - Square Matrix

L- Lower Triangle Matrix(has all the elements above the main diagonal as zero)

U- Upper Triangle Matrix(has all the elements below the main diagonal as zero)

LU Decomposition with Partial Pivoting

LU decomposition is found using an iterative numerical process and can fail for those matrices that cannot be decomposed or decomposed easily. A variation of this decomposition that is numerically more stable to solve in practice is called the LUP decomposition, or the LU decomposition with partial pivoting

$$A = L \cdot U \cdot P$$

The rows of the parent matrix are re-ordered to simplify the decomposition process and the additional P matrix specifies a way to permute the result or return the result to the original order.

Application

1. Solving of systems of linear equations (finding the coefficients in a linear regression).
2. Calculating the determinant and inverse of a matrix.

QR Decomposition

1. Used for n x m matrices.

$$A = Q \cdot R$$

Q a matrix with the size $m \times m$

R - upper triangle matrix with the size $m \times n$

Applications

1. Same as LU Decomposition but can be applied on non square matrices.

Cholesky Decomposition

1. Used for Positive Definite Matrices(square symmetric matrices where all values are greater than zero).

$$A = L \cdot L^T$$

L is the lower triangular matrix

L^T is the transpose of L

Or, It can also be written as the product of the upper triangular matrix.

$$A = U^T \cdot U$$

U is the upper triangular matrix

Advantage

1. While decomposing symmetric matrices, the Cholesky decomposition is nearly twice as efficient as the LU decomposition and should be preferred in these cases.

Application

1. Solving of systems of linear equations (finding the coefficients in a linear regression).
2. Calculating the determinant and inverse of a matrix.

Eigen Decomposition

1. Decompose a square matrix into a set of eigenvectors and eigenvalues.
2. A vector is an eigenvector of a matrix if it satisfies the eigenvalue equation:

$$A \cdot v = \lambda v$$

A is the parent square matrix that we are decomposing, v is the eigenvector of the matrix, and λ represents the eigenvalue scalar

To understand in simple terms.

Generally, almost all vectors change direction, when they are multiplied by A. Certain exceptional vectors v are in the same direction as Av. Those are the **eigenvectors**.

Multiply an eigenvector by A, and the vector Av is the number λ times the original v. The eigenvalue λ tells whether the special vector v is stretched or shrunk or reversed or left unchanged — when it is multiplied by A.

3. A matrix could have one eigenvector and eigenvalue for each dimension of the parent matrix.
4. The parent matrix can be shown to be a product of the eigenvectors and eigenvalues.

$$A = Q \cdot \Lambda \cdot Q^T$$

Q is a matrix comprised of the eigenvectors

Λ is the uppercase Greek letter lambda and is the diagonal matrix comprised of the eigenvalues

Q^T is the transpose of the matrix comprised of the eigenvectors

5. Eigenvectors are unit vectors, which means that their length or magnitude is equal to 1.0.
6. Eigenvalues are coefficients applied to eigenvectors that give the vectors their length or magnitude.
7. A matrix that has only positive eigenvalues is referred to as a positive definite matrix, whereas if the eigenvalues are all negative, it is referred to as a negative definite matrix.

Application

1. To calculate the principal components of a matrix that can be used to reduce the dimensionality of data in machine learning.
2. Gives valuable insights into the properties of the matrix that makes certain matrix calculations easier like computing the power of the matrix.

Singular Value Decomposition

1. Decompose $n \times m$ matrix into its constituent parts.

$$A = U \Sigma V^T$$

A – $n \times m$ matrix to decompose

U – $m \times m$ matrix - Left-singular vectors of A

Σ – $m \times n$ diagonal matrix called Singular Values

V^T – Transpose of $n \times n$ matrix - Right-singular vectors of A

Application

1. Data reduction method in machine learning.
2. Least squares linear regression.
3. Image compression
4. Denoising data

Pseudoinverse

1. It is the generalisation of the matrix inverse for square matrices to rectangular matrices where number of rows and columns are not equal.
2. It is also called the Moore-Penrose Inverse.

$$A^+ = V \cdot D^+ \cdot U^T$$

A^+ is the pseudoinverse

D^+ is the pseudoinverse of the diagonal matrix Σ

U^T is the transpose of U

3. We can get U and V from the SVD operation.

$$A = U \Sigma V^T$$

4. The D^+ can be calculated by creating a diagonal matrix from Σ , calculating the reciprocal of each non-zero element in Σ , and taking the transpose if the original matrix was rectangular.

Dimensionality Reduction using SVD

1. Decompose a matrix that results in a matrix with a lower rank that is said to approximate the original matrix.
2. We use SVD operation on the original data and select the top k largest singular values in Σ . These columns can be selected from Σ and the rows selected from V^T . This is called Truncated SVD.

$$B = U \Sigma_k V_k^T$$

3. In natural language processing, this approach can be used on matrices of word occurrences or word frequencies in documents and is called Latent Semantic Analysis or Latent Semantic Indexing. We calculate the dense summary of the matrix as:

$$T = U \Sigma_k$$

4. To transform the original matrix A or other similar matrices at the inference time we can do:

$$T = A \cdot V_k^T$$

PCA

References

1. [Basics of Linear Algebra for Machine Learning](#)