Chapter 14
Rank of a Matix and Singular Value Decomposition

1 Rank of a Matrix

Rank of a matrix is the number of linearly independent columns of a matrix
For example,
Consider this matrix

$$\begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}$$

Can the first column $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ be written as a linear combination of its previous columns?
Since there are no previous columns, this column is linearly independent of its previous columns, hence we can put a $\times$ over it.

$$\begin{array}{cc} \times & \\ \end{array}$$

$$\begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}$$

Can the second column $\begin{bmatrix} 1 \\ 2 \end{bmatrix}$ be written as a linear combination of its previous columns?
No, we can not write $c\begin{bmatrix} 1 \\ 1 \end{bmatrix}=\begin{bmatrix} 1 \\ 2 \end{bmatrix}$
Hence, this column is also linearly independent. We can put a $\times$ over it also.

$$\begin{array}{cc} \times & \times \\ \end{array}$$

$$\begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}$$

We are done with all the columns, now to calculate the rank of this matrix we just need to calculate number of $\times$ over this matrix.
Hence, Rank$\left(\begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}\right)$=2

More examples

1.

$$\times$$

$$Rank\left(\begin{bmatrix}1 & 2\\ 1 & 2\end{bmatrix}\right) = 1$$

2.

$$Rank\left(\begin{bmatrix}0 & 0\\ 0 & 0\end{bmatrix}\right) = 0$$

Now, let's consider a bigger matrix

$$\begin{bmatrix}1 & 1 & 2 & 4 & 2\\ 2 & 1 & 3 & 5 & 4\\ 1 & 1 & 2 & 4 & 2\\ 0 & 1 & 1 & 3 & 0\end{bmatrix} = \begin{bmatrix}\uparrow & \uparrow & \uparrow & \uparrow & \uparrow & \\ a_1 & a_2 & a_3 & a_4 & a_5 & \\ \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \end{bmatrix}$$

- $a_1 = \begin{bmatrix}1\\2\\1\\0\end{bmatrix}$ can't be written as a linear combination of it's previous columns
- $a_2 = \begin{bmatrix}1\\1\\1\\1\end{bmatrix}$ can't be written as a linear combination of it's previous columns
- $a_3 = \begin{bmatrix}2\\3\\2\\1\end{bmatrix}$ can be written as a linear combination of it's previous columns

$$\begin{bmatrix}1\\2\\1\\0\end{bmatrix} + \begin{bmatrix}1\\1\\1\\1\end{bmatrix} = \begin{bmatrix}2\\3\\2\\1\end{bmatrix}$$
$$a_1 + a_2 = a_3$$

- $a_4 = \begin{bmatrix}4\\5\\4\\3\end{bmatrix}$ can be written as a linear combination of it's previous columns

$$1\begin{bmatrix}1\\2\\1\\0\end{bmatrix} + 3\begin{bmatrix}1\\1\\1\\1\end{bmatrix} + 0\begin{bmatrix}2\\3\\2\\1\end{bmatrix} = \begin{bmatrix}4\\5\\4\\3\end{bmatrix}$$

$$1a_1 + 3a_2 + 0.a_3 = a_4$$

- $a_5 = \begin{bmatrix} 2 \\ 4 \\ 2 \\ 0 \end{bmatrix}$ can be written as a linear combination of it's previous columns

$$2 \begin{bmatrix} 1 \\ 2 \\ 1 \\ 0 \end{bmatrix} + 0 \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} + 0 \begin{bmatrix} 2 \\ 3 \\ 2 \\ 1 \end{bmatrix} + 0 \begin{bmatrix} 4 \\ 5 \\ 4 \\ 3 \end{bmatrix} = \begin{bmatrix} 2 \\ 4 \\ 2 \\ 0 \end{bmatrix}$$

$$2a_1 + 0a_2 + 0a_3 + 0a_4 = a_5$$

Hence

$$\begin{array}{c} \begin{bmatrix} \times & \times & . & . & . \end{bmatrix} \\ Rank \left(\begin{bmatrix} 1 & 1 & 2 & 4 & 2 \\ 2 & 1 & 3 & 5 & 4 \\ 1 & 1 & 2 & 4 & 2 \\ 0 & 1 & 1 & 3 & 0 \end{bmatrix} \right) = 2 \end{array}$$

What does this means?
This means we can write all the columns of this matrix using a linear combination of only the 1st and 2nd column i.e., a_1 and a_2.
This means a_1 and a_2 are the linearly independent basis vectors for $a_3, a_4 \& a_5$ this means

- a_1=$1a_1$+$0a_2$
- a_2=$0a_1$+$1a_2$
- a_3=$1a_1$+$1a_2$
- a_4=$1a_1$+$3a_2$
- a_5=$2a_1$+$0a_2$

$$\Longrightarrow \begin{bmatrix} \uparrow & \uparrow & \uparrow & \uparrow & \uparrow & \\ a_1 & a_2 & a_3 & a_4 & a_5 & \\ \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \end{bmatrix} = \begin{bmatrix} \uparrow & \uparrow \\ a_1 & a_2 \\ \downarrow & \downarrow \end{bmatrix} . \begin{bmatrix} 1 & 0 & 1 & 1 & 2 \\ 0 & 1 & 1 & 3 & 0 \end{bmatrix}$$

- Hence, there exist $Rank(A)$ basis vectors for a matrix A.
- If M is a square matrix of dimensions $n \times n$ then there exist n eigenvectors which can act as a basis for this matrix.

  But what if $Rank(M) < n$

  is this conflicting with our result that "there exist $Rank(A)$ basis vectors for a matrix A"? Not really.

  If $Rank(M) < n$ then there would still be n eigenvectors but only $Rank(M)$ non-zero eigenvalues.

  These $n - Rank(M)$ zero eigenvalues when multiplied with their corressponding eigenvectors will make them zero vectors.

  for example $Rank\left(\begin{bmatrix}1 & 2\\ 1 & 2\end{bmatrix}\right) = 1$ but number of eigenvectors=2
  Let v_1, v_2 be eigenvectors of this matrix and λ_1, λ_2 be eigenvalues of this matrix
  $v_1 = \begin{bmatrix}-2\\ 1\end{bmatrix}$ and $v_2 = \begin{bmatrix}1\\ 1\end{bmatrix}$
  But $\lambda_1 = 0$ and $\lambda_2 = 3$

  This means there is only 1 non-zero basis vector of this matrix with $Rank = 1$
- Hence, the correct result is: there exist $Rank(A)$ non-zero basis vectors for a matrix A.

2 Problem

If A is a square matrix then we can write

$$
\begin{aligned}
Av_1 &= \lambda_1 v_1 \\
Av_2 &= \lambda_2 v_2 \\
&\vdots \\
Av_n &= \lambda_n v_n
\end{aligned}
$$

Where $v_1, v_2, \ldots, v_n$ are the eigenvectors of A and $\lambda_1, \lambda_2, \ldots, \lambda_n$ are the eigenvalues of A.
If $v_1, v_2, \ldots, v_n$ are taken as basis, then we can write any vector $x \in \mathbb{R}^n$ as a linear combination of these basis eigenvectors.

$$x = \alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_n v_n = \sum_{i=1}^{n} \alpha_i v_i$$

So, what would be matrix vector product Ax be?

$$Ax = \sum_{i=1}^{n} \alpha_i . A . v_i = \sum_{i=1}^{n} \alpha_i \lambda_i v_i$$

A matrix vector product became a scalar vector product...!!!
That's one of the advantage of having a square matrix is that we can have eigenvectors and eigenvalues and convert that matrix's operations into something simpler.

Can we have eigenvectors for a non-square i.e., a rectangular matrix?
In other words, is this possible

$$A_{(\text{m x n})} x_{(\text{n x 1})} = x_{(\text{n x 1})}?$$

No!, Why?
Because

$$\mathbb{R}^{(\text{m x n})} . \mathbb{R}^{(\text{n x 1})} = \mathbb{R}^{(\text{m x 1})}$$

$$\begin{bmatrix} \quad (\text{m x n}) Matrix \quad \end{bmatrix} . \begin{bmatrix} \\ (\text{n x 1}) \\ Vector \\ \\ \end{bmatrix} \implies \begin{bmatrix} (\text{m x 1}) \\ Vector \end{bmatrix}$$

Since, any vector cannot remain of the same dimensions after rectangular matrix transformation.
Hence eigenvectors don't exist for rectangular matrices.
Can we not have something that can change a matrix operation into some scalar operations for a rectangular matrices then?

3 Setup

Ok, so we can think of a rectangular matrix $\mathbb{R}^{\text{m x n}}$ as a function that takes a $\mathbb{R}^n$ matrix and outputs a $\mathbb{R}^m$ matrix.
If $(v_1, u_1), (v_2, u_2), \ldots, (v_k, u_k)$ are pairs of vectors such that
$v_i \in \mathbb{R}^n$ and $u_i \in \mathbb{R}^m$
then we hope to write.

$$Av_i = \sigma_i u_i$$

Where, $A \in \mathbb{R}^{\text{m x n}}$
And if this is true and if a σ_i exist
And if we assume that $v_1, v_2, \ldots, v_k$ are orthonormal and thus form a basis V in $\mathbb{R}^n$ then we can write any $x \in \mathbb{R}^n$ as a linear combination of these $v_1, v_2, \ldots, v_k$ basis vectors

$$x = \alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_k v_k = \sum_{i=1}^{k} \alpha_i v_i$$

But do you see something, x is an n dimensional vector and we are trying to represent it using k basis's which means k dimensions, why?
Recall that if $M \in \mathbb{R}^n$ is a square matrix of dimensions n x n then, there will be n eigenvectors for it.
And we know eigenvectors are linearly independent and can thus form a basis for any vector $x \in \mathbb{R}^n$.
So, we can say a square matrix of dimensions $n \times n$ can always have n basis vectors.
But we cannot say the same for non-square aka rectangular matrices.
There can only be $Rank(M)$ non-zero basis vectors for a non-square matrix.

Hence the dimensions of x will have k non-zero basis vectors, where $k = Rank(A)$
Therefore,

$$x = \alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_k v_k = \sum_{i=1}^{k} \alpha_i v_i$$

4 Finding the Reduced form of $A_{m \times n}$

And If $Av_i = \sigma_i u_i$ was possible then we can write

$$A_{m \times n}.V_{n \times k} = U_{m \times k}.\Sigma_{k \times k}$$

$$A_{m \times n}. \begin{bmatrix} \uparrow & \uparrow & \dots & \uparrow \\ v_1 & v_2 & \dots & v_k \\ \downarrow & \downarrow & \dots & \downarrow \end{bmatrix} = \begin{bmatrix} \uparrow & \uparrow & \dots & \uparrow \\ u_1 & u_2 & \dots & u_k \\ \downarrow & \downarrow & \dots & \downarrow \end{bmatrix} . \begin{bmatrix} \sigma_1 & 0 & \dots & 0 \\ 0 & \sigma_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \sigma_k \end{bmatrix}$$

Here,

- V is a matrix of input basis vectors
- U is a matrix of ouput basis vectors

Since, we have only k orthogonal basis vectors for V and U and there are $n - k$ basis vectors remaning we can find these remaining vectors using Gram Schmidt orthogonalisation process.
After getting all n orthogonal basis vectors

$$A_{m \times n}.V_{n \times n} = U_{m \times n}.\Sigma_{n \times n}$$

$$A_{m \times n}. \begin{bmatrix} \uparrow & \uparrow & \dots & \uparrow \\ v_1 & v_2 & \dots & v_n \\ \downarrow & \downarrow & \dots & \downarrow \end{bmatrix} = \begin{bmatrix} \uparrow & \uparrow & \dots & \uparrow \\ u_1 & u_2 & \dots & u_n \\ \downarrow & \downarrow & \dots & \downarrow \end{bmatrix} . \begin{bmatrix} \sigma_1 & 0 & \dots & 0 \\ 0 & \sigma_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \sigma_n \end{bmatrix}$$

Then we can say,
Since, V and U are orthogonal matrices, this means
$V^T V = \mathbb{I}$ [Identity matrix]
$U^T U = \mathbb{I}$ [Identity matrix]

But we also know for any matrix M, $M^{-1}M = \mathbb{I}$
This means
$V^T = V^{-1}$ if V is an orthogonal matrix
$U^T = U^{-1}$ if U is an orthogonal matrix

And if,

$$A.V = U.\Sigma$$

Then,

$$U^{-1}AV = \Sigma = U^T AV \ [Diagonalisation\ of\ A]$$

$$A = U.\Sigma.V^{-1} = U.\Sigma.V^T \ [Singular\ Value\ Decomposition\ of\ A]$$

This is called Singular Value Decomposition, as we are decomposing a non-square matrix A into simpler vector matrices and its singular values which allows us for simpler operations on A.

Singular Values

Singular values of a matrix M are the positive square roots of the eigenvalues of $M^T M$

Suppose $V, U \& \Sigma$ exist, then

$$A^T_{n\times m} A_{m\times n} = M_{n\times n} = (U.\Sigma.V^T)^T.(U.\Sigma.V^T) = V.\Sigma^T.U^T.U.\Sigma.V^T = V.\Sigma^T.\mathbb{I}.\Sigma.V^T = V.\Sigma^T.\Sigma.V^T$$

But since Σ is a diagonal matrix $\Sigma^T\Sigma = \Sigma^2$
Hence

$$A^T A = V.\Sigma^T.\Sigma.V^T = V.\Sigma^2.V^T$$

$$A^T A = V.\Sigma^2.V^T$$

Similarly

$$AA^T = U.\Sigma^2.U^T$$

If we Recall,
If S is a square symmetric matrix, E is a matrix of orthonormal eigenvectors of S and Λ is a diagonal matrix of eigenvalues of S

$$S = E\Lambda E^T$$

Is the Eigenvalue Decomposition of S

In our situation also

Singular Value Decomposition $\iff$ Eigenvalue Decomposition

For
$$A^T A = V.\Sigma^2.V^T$$

- $A^T A$ is a square symmetric matrix
- V is the matrix of eigenvectors of $A^T A$
- V is also called the right singular vectors of matrix A
- $\Sigma^2 = \Lambda$ is the diagonal matrix of eigenvalues of $A^T A$.
- Σ^2 These eigenvalues are also the squares of singular values of A

Singular Value Decomposition $\iff$ Eigenvalue Decomposition

and for

$$AA^T = U.\Sigma^2.U^T$$

- AA^T is a square symmetric matrix
- U is the matrix of eigenvectors of AA^T
- U is also called the left singular vectors of matrix A
- $\Sigma^2 = \Lambda$ is the diagonal matrix of eigenvalues of AA^T.
- Σ^2 These eigenvalues are also the squares of singular values of A