Matrices

Concepts revisited

NULL SPACE

0.1 Trace of a matrix

The **trace** of an $n \times n$ matrix $\boldsymbol{A}$ is defined as the sum of its elements along its **principal diagonal**.

$$trace(\boldsymbol{A}) = a_{11} + a_{22} + \cdots + a_{nn} \tag{1}$$

Further, if $\boldsymbol{A}$ is an $m \times n$ matrix and $\boldsymbol{B}$ is an $n \times m$ matrix then if we take their matrix product $\boldsymbol{AB}$ $(m \times m)$, its trace would be given as follows:

$$trace(\boldsymbol{AB}) = \sum_{j=1}^{n} a_{1j}b_{j1} + \sum_{j=1}^{n} a_{2j}b_{j2} + \cdots + \sum_{j=1}^{n} a_{nj}b_{jn} \tag{2}$$

The above can be thought of as follows :- Since trace is the sum of the diagonal elements, when we multiply two matrices, the first diagonal element is nothing but the result of the dot product between the first row of martix $\boldsymbol{A}$ with the first column of matrix $\boldsymbol{B}$. The second diagonal element would be the dot product of the second row of matrix $\boldsymbol{A}$ with the second column of matrix $\boldsymbol{B}$ and so forth. That is why the Trace in this case is a summation of summation elements. Remember some properties of the **trace operator**:

- $trace(\boldsymbol{AB}) = trace(\boldsymbol{BA})$
- $trace(\boldsymbol{A} + \boldsymbol{B}) = trace(\boldsymbol{A}) + trace(\boldsymbol{B})$. Only if $\boldsymbol{A}$ and $\boldsymbol{B}$ are $n \times n$.
- $trace(\lambda \boldsymbol{A}) = \sum_{j=1}^{n} \lambda a_{ii} = \lambda \sum_{j=1}^{n} a_{ii} = \lambda \times trace(\boldsymbol{A})$. Only if $\boldsymbol{A}$ is $n \times n$.

0.2 Partitioned matrices

A **partitioned matrix** is a matrix whose individual elements are also matrices. Consider the following matrix:

$$\boldsymbol{A} = \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \end{bmatrix} \tag{3}$$

This could be rewritten as follows:

$$\boldsymbol{A} = \begin{bmatrix} \boldsymbol{A_1} & \boldsymbol{A_2} \\ \boldsymbol{a}_1^T & \boldsymbol{a}_2^T \end{bmatrix} \tag{4}$$

Where we have the following elemental matrices defined as follows:

$$\boldsymbol{A_1} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}, \ \boldsymbol{A_2} = \begin{bmatrix} a_{13} & a_{14} \\ a_{23} & a_{24} \end{bmatrix} \tag{5}$$

$$\boldsymbol{a}_1^T = [a_{31}, a_{32}], \ \boldsymbol{a}_2^T = [a_{33}, a_{34}] \tag{6}$$

These partitioned matrices can be subject to usual mathematical matrix operations like addition, subtraction and multiplication, provided the corresponding matrix dimensions are in sync.

0.3 Determinants

The determinant of a simple 2×2 matrix is pretty simple to compute:

$$|\boldsymbol{A}| = a_{11}a_{22} - a_{12}a_{21} \tag{7}$$

When it comes to computing the determinant of a $n \times n$ matrix, we first let $\boldsymbol{A_{ij}}$ denote the $(n-1) \times (n-1)$ matrix formed by deleting the row i and column j of the original matrix $\boldsymbol{A}$. The determinant is then given by the general formula:

$$|\boldsymbol{A}| = \sum_{j=1}^{n} (-1)^{j+1} a_{1j} |\boldsymbol{A_{1j}}| \tag{8}$$

To demonstrate more clearly, here is how the determinant of a 3×3 matrix is computed:

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{11} \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{12} \begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} + a_{13} \begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix} \tag{9}$$

Some general properties about determinants:

- Property 1:
$$|\boldsymbol{AB}| = |\boldsymbol{A}| \cdot |\boldsymbol{B}| \tag{10}$$
- Property 2: The determinant of a **lower triangular matrix** and an **upper triangular matrix** is nothing but the product of the principal diagonal elemnts.
- Property 3:
$$|\boldsymbol{A}^T| = |\boldsymbol{A}| \tag{11}$$

0.4 Inverse

Before calculating the inverse of a matrix, we need to calculate the **adjoint of a matrix**. And before we compute the adjoint, we must again recall that if A is an $n \times n$ matrix, then $\boldsymbol{A_{ji}}$ represents the $(n-1) \times (n-1)$ matrix that is formed when we delete row j and column i from $\boldsymbol{A}$. Now the adjoint of matrix $\boldsymbol{A}$ is an $n \times n$ matrix whose $(i,j)^{th}$ element is given by:

$$(-1)^{i+j}|\boldsymbol{A_{ji}}| \tag{12}$$

Now we note that if the determinant of an $n \times n$ matrix $\boldsymbol{A}$ is not equal to 0, then its inverse exists and is found by dividing the adjoint of the matrix by its determinant:

$$\boldsymbol{A^{-1}} = \frac{1}{|\boldsymbol{A}|} \cdot [(-1)^{i+j}|\boldsymbol{A_{ji}}|] \tag{13}$$

We further note that a matrix whose inverse exists is called **nonsingular** and a matrix whose determinant is 0 is called **singular** and has no inverse. We note the additional property that:

$$\boldsymbol{A} \times \boldsymbol{A^{-1}} = \boldsymbol{I_n} \tag{14}$$

$$|\boldsymbol{A^{-1}}| = \frac{1}{|\boldsymbol{A}|} \tag{15}$$

Lastly we note the following additional properties for inverses:

- $(\boldsymbol{A^{-1}})^T = (\boldsymbol{A^T})^{-1}$
- $[\alpha \boldsymbol{A}]^{-1} = \alpha^{-1}\boldsymbol{A^{-1}}$
- $[\boldsymbol{AB}]^{-1} = \boldsymbol{B^{-1}A^{-1}}$
- $[\boldsymbol{ABC}]^{-1} = \boldsymbol{C^{-1}B^{-1}A^{-1}}$

0.5 Linear dependence

Let $\boldsymbol{x_1}, \boldsymbol{x_2}, \cdots, \boldsymbol{x_k}$ be k different $(n \times 1)$ vectors. These vectors are said to be **linearly dependent** is there exist a set of k scalars $(c_1, \cdots, c_k)$ such that not all of them are zero and the following holds:

$$c_1\boldsymbol{x_1} + c_2\boldsymbol{x_2} + \cdots + c_k\boldsymbol{x_k} = 0 \tag{16}$$

Note that if no such non zero set of scalars exist, then the vectors are said to be **linearly independent**. Now suppose we collect all these vectors $\boldsymbol{x_1}, \cdots, \boldsymbol{x_k}$ into an $n \times k$ matrix $\boldsymbol{T}$ such that:

$$\boldsymbol{T} = \begin{bmatrix} \boldsymbol{x_1} & \cdots & \boldsymbol{x_k} \end{bmatrix} \tag{17}$$

Now further if we suppose k to be equal to n, that is we take a collection n vectors, each of $n \times 1$ dimension, then in that case $\boldsymbol{T}$ would become an $n \times n$ matrix. And further if $\boldsymbol{x_1}, \cdots, \boldsymbol{x_n}$ turn out to be linearly dependent, then we find that the determinant of matrix $\boldsymbol{T}$ is zero. Suppose that $\boldsymbol{x_1}$ is one such vector that has a nonzero scalar weight c_1. Then linear dependence would imply that $\boldsymbol{x_1}$ can be written in terms of a weighted sum of the other vectors as follows:

$$\boldsymbol{x_1} = -(c_2/c_1)\boldsymbol{x_2} - (c_3/c_1)\boldsymbol{x_3} - \cdots - (c_n/c_1)\boldsymbol{x_n} \tag{18}$$

0.6 Eigenvalues and Eigenvectors

Suppose that an $n \times n$ matrix $\boldsymbol{A}$, a nonzero $n \times 1$ vector $\boldsymbol{x}$ and a scalar λ are related by the following equation:

$$\boldsymbol{A}\boldsymbol{x} = \lambda\boldsymbol{x} \tag{19}$$

Then we say that $\boldsymbol{x}$ is the **Eigenvector** of $\boldsymbol{A}$ and λ is its associated **Eigenvalue**. The above equation can be further rewritten as:

$$\boldsymbol{A}\boldsymbol{x} - \lambda\boldsymbol{I_n}\boldsymbol{x} = 0 \tag{20}$$

$$(\boldsymbol{A} - \lambda\boldsymbol{I_n})\boldsymbol{x} = 0 \tag{21}$$

Now we say that provided $(\boldsymbol{A} - \lambda\boldsymbol{I_n})$ is nonsingular and its inverse exists, then a nonzero vector $\boldsymbol{x}$ would satisfy the above equation. Further we say that Eigenvalue is a numbe such that:

$$|\boldsymbol{A} - \lambda\boldsymbol{I_n}| = 0 \tag{22}$$

Now suppose that matrix $\boldsymbol{A}$ is a lower or upper triangular matrix, in this case it is obvious that even $(\boldsymbol{A} - \lambda\boldsymbol{I_n})$ would be a lower or upper triangular matrix as well. As we saw earlier, the determinant of such a matrix is just product of its principal diagonal terms:

$$|\boldsymbol{A} - \lambda\boldsymbol{I_n}| = (a_{11} - \lambda)(a_{22} - \lambda)\cdots(a_{nn} - \lambda) \tag{23}$$

We can see that this expression would be zero if the eigenvalue would take on a value equal to any of the diagonal terms of the matrix $\boldsymbol{A}$. And that is indeed the case - the eigenvalues of a triangular matrix are just the values of $\boldsymbol{A}$ along the principal diagonal.

0.6.1 Linear independence of Eigenvectors

A common result often encountered is that if the eigenvalues $(\lambda_1, \cdots, \lambda_n)$ are **distinct** then the associated eigenvectors $(\boldsymbol{x_1}, \cdots, \boldsymbol{x_n})$ are all linearly independent. Let us consider a simple case to illustrate this concept by taking $n = 2$. Consider any two number c_1 and c_2 such that:

$$c_1\boldsymbol{x_1} + c_2\boldsymbol{x_2} = 0 \tag{24}$$

Premultiplying both sides by A we get:

$$c_1\boldsymbol{A}\boldsymbol{x_1} + c_2\boldsymbol{A}\boldsymbol{x_2} = c_1\lambda_1\boldsymbol{x_1} + c_2\lambda_2\boldsymbol{x_2} = 0 \tag{25}$$

If we multiply equation 24 with λ_1 and subtract it from equation 25 we would get the following:

$$c_2(\lambda_2 - \lambda_1)\boldsymbol{x_2} = 0 \tag{26}$$

Now we note that since $\boldsymbol{x_2}$ is an eigenvector of $\boldsymbol{A}$ it cannot be zero. Even $\lambda_2 - \lambda_1$ can't be zero since $\lambda_2 \neq \lambda_1$. So it follows that $c_2 = 0$. By similar arguments we can prove that even $c_1 = 0$, which would inturn prove that $\boldsymbol{x_2}$ and $\boldsymbol{x_1}$ are infact, linearly independent.

0.6.2 Eigen Decomposition

Suppose an $n \times n$ matrix $\boldsymbol{A}$ has n distinct eigenvalues $(\lambda_1, \cdots, \lambda_n)$. If we collect all these values into a diagonal matrix $\boldsymbol{\Lambda}$ we would get:

$$\Lambda = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix} \tag{27}$$

Further, we collect all the associated eigenvectors $(\boldsymbol{x_1}, \cdots, \boldsymbol{x_n})$ in an $n \times n$ matrix $\boldsymbol{T}$ as follows:

$$\boldsymbol{T} = \begin{bmatrix} \boldsymbol{x_1} & \boldsymbol{x_2} & \cdots & \boldsymbol{x_n} \end{bmatrix} \tag{28}$$

Premultiplying matrix $\boldsymbol{A}$ would then give us:

$$\boldsymbol{AT} = \begin{bmatrix} \boldsymbol{Ax_1} & \boldsymbol{Ax_2} & \cdots & \boldsymbol{Ax_n} \end{bmatrix} \tag{29}$$

Now since the $\boldsymbol{x}$ vectors are eigenvectors and that $\boldsymbol{Ax} = \lambda\boldsymbol{x}$, we can rewrite the above equation as follows:

$$\boldsymbol{AT} = \begin{bmatrix} \lambda_1\boldsymbol{x_1} & \lambda_2\boldsymbol{x_2} & \cdots & \lambda_n\boldsymbol{x_n} \end{bmatrix} \tag{30}$$

The RHS of the above equation can be thought of as the matrix product of the matrix of eigenvectors and the diagonal eigenvalue matrix:

$$\begin{bmatrix} \lambda_1\mathbf{1} & \cdots & \lambda_n\boldsymbol{x_n} \end{bmatrix} = \begin{bmatrix} \boldsymbol{x_1} & \cdots & \boldsymbol{x_n} \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix} = \boldsymbol{T\Lambda} \tag{31}$$

So we can finally write:

$$\boldsymbol{AT} = \boldsymbol{T\Lambda} \tag{32}$$

Now we know that since the eigenvectors are linearly independent, the matrix $\boldsymbol{T}$ which is just a collection of eigenvectors, has to be nonsingular. So we know that its inverse exists. With this, we can finally decompose $\boldsymbol{A}$ as:

$$\boldsymbol{A} = \boldsymbol{T\Lambda T^{-1}} \tag{33}$$

0.6.3 The Jordan decomposition

Note that the previous result was only possible because the $n \times n$ matrix $\boldsymbol{A}$ was assumed to have n distinct eigenvalues and consequently, n linearly independent eigenvectors. If however, we had the case wherein $\boldsymbol{A}$ had $s \leq n$ linearly independent eigenvectors, we would use the **Jordan decomposition** on $\boldsymbol{A}$. In this case, we say that for matrix $\boldsymbol{A}$, there exists a nonsingular matrix $\boldsymbol{M}$ such that:

$$\boldsymbol{A} = \boldsymbol{M}\boldsymbol{J}\boldsymbol{M}^{-1} \tag{34}$$

Where $\boldsymbol{J}$ is an $n \times n$ matrix of the form:

$$\boldsymbol{J} = \begin{bmatrix} \boldsymbol{J_1} & \boldsymbol{0} & \cdots & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{J_2} & \cdots & \boldsymbol{0} \\ \vdots & \vdots & \cdots & \vdots \\ \boldsymbol{0} & \boldsymbol{0} & \cdots & \boldsymbol{J_s} \end{bmatrix} \tag{35}$$

Where we have each $\boldsymbol{J_i}$ submatrix containing the repeated eigenvalue along its principal diagonal, with the element 1 repeated above the diagonal elements:

$$\boldsymbol{J_i} = \begin{bmatrix} \lambda_i & 1 & 0 & \cdots & 0 \\ 0 & \lambda_i & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \cdots & 0 \\ 0 & 0 & 0 & \cdots & \lambda_i \end{bmatrix} \tag{36}$$

0.7 Positive definite matrices

An $n \times n$ real symmetric matrix $\boldsymbol{A}$ is said to be **positive semidefinite** is for any real $n \times 1$ vector $\boldsymbol{x}$, we have:

$$\boldsymbol{x^T}\boldsymbol{A}\boldsymbol{x} \geq 0 \tag{37}$$

A stronger statement than this would be that of **positive definiteness** which states:

$$\boldsymbol{x^T}\boldsymbol{A}\boldsymbol{x} > 0 \tag{38}$$

Now let λ be the eigenvalue of matrix $\boldsymbol{A}$ associated with eigenvector $\boldsymbol{x}$:

$$\boldsymbol{A}\boldsymbol{x} = \lambda\boldsymbol{x} \tag{39}$$

If we premultiply the above equation with $\boldsymbol{x^T}$ we get:

$$\boldsymbol{x^T}\boldsymbol{A}\boldsymbol{x} = \lambda\boldsymbol{x^T}\boldsymbol{x} \tag{40}$$

Now we know that the eigenvectors cannot be a zero vector, therefore $\boldsymbol{x^T}x > 0$. From this it would naturally follow that the eigenvalues of a positive semidefinite matrix $\boldsymbol{A}$ must be greater than or equal to 0. And for a positive definite matrix $\boldsymbol{A}$ all the eigenvalues are strictly greater than 0. Note that since the determinant of $\boldsymbol{A}$ is the **product of its eigenvalues**, the determinant of a positive definite matrix $\boldsymbol{A}$ is strictly positive.

References

[1] Hamilton - Time Series Analysis (appendix)