The Trace of a Matrix

The trace of a square matrix $\underset{n\times n}{\boldsymbol{A}} = [a_{ij}]$ is

$$trace(\boldsymbol{A}) = tr(\boldsymbol{A}) = \sum_{i=1}^{n} a_{ii}.$$

For example,

$$tr\left(\begin{bmatrix} 5 & 3 & 5 \\ 4 & -1 & 2 \\ -3 & 8 & 7 \end{bmatrix}\right) = 5 - 1 + 7 = 11.$$

Some Simple Facts about Trace

Suppose $k, k_1, \ldots, k_m \in \mathbb{R}$ and $\boldsymbol{A}, \boldsymbol{A}_1, \ldots, \boldsymbol{A}_m$ are each $n \times n$ matrices. Then

1. $tr(\boldsymbol{A}) = tr(\boldsymbol{A}')$
2. $tr(k\boldsymbol{A}) = k \cdot tr(\boldsymbol{A})$
3. $tr(\boldsymbol{A}_1 + \boldsymbol{A}_2) = tr(\boldsymbol{A}_1) + tr(\boldsymbol{A}_2)$
4. $tr(\sum_{i=1}^m k_i\boldsymbol{A}_i) = \sum_{i=1}^m k_i \cdot tr(\boldsymbol{A}_i)$

Result A.17:

(a) $tr(\boldsymbol{AB}) = tr(\boldsymbol{BA})$. This is known as the cyclic property of the trace.

(b) If $\underset{m\times n}{\boldsymbol{A}} = [a_{ij}]$, then

$$tr(\boldsymbol{A}'\boldsymbol{A}) = \sum_{i=1}^{m}\sum_{j=1}^{n} a_{ij}^2.$$

Proof of Result A.17: HW problem. □

Suppose $\boldsymbol{A}$ is an $m \times n$ matrix of rank r. Prove that there exist matrices $\underset{m\times r}{\boldsymbol{B}}$ and $\underset{r\times n}{\boldsymbol{C}}$ such that

$$\boldsymbol{A} = \boldsymbol{BC} \text{ and } rank(\boldsymbol{B}) = rank(\boldsymbol{C}) = r.$$

Proof:

Let $\boldsymbol{B} = [\boldsymbol{b}_1, \ldots, \boldsymbol{b}_r]$ where $\boldsymbol{b}_1, \ldots, \boldsymbol{b}_r$ form a basis for $\mathcal{C}(\boldsymbol{A})$.

Because $\boldsymbol{b}_1, \ldots, \boldsymbol{b}_r$ form a basis, they are LI so that $rank(\boldsymbol{B}) = r$.

Let $\boldsymbol{c}_j$ be the vector of the coefficients of the linear combination of $\boldsymbol{b}_1, \ldots, \boldsymbol{b}_r$ that gives the jth column of $\boldsymbol{A}$.

Then $\boldsymbol{A} = \boldsymbol{B}\boldsymbol{C}$, where $\boldsymbol{C} = [\boldsymbol{c}_1, \ldots, \boldsymbol{c}_n]$.

Finally, note that

$$r = rank(\boldsymbol{A}) = rank(\boldsymbol{B}\boldsymbol{C}) \leq rank(\boldsymbol{C}) \leq r \Rightarrow rank(\boldsymbol{C}) = r. \quad \square$$

Suppose $\boldsymbol{A}$ is an $n \times n$ matrix such that $\boldsymbol{AA} = k\boldsymbol{A}$ for some $k \in \mathbb{R}$.
Prove that $tr(\boldsymbol{A}) = k \cdot rank(\boldsymbol{A})$.

(Note that this result implies the trace of an idempotent matrix is equal to its rank.)

Proof:

Let $r = rank(\boldsymbol{A})$. Let $\underset{n\times r}{\boldsymbol{B}}$ and $\underset{r\times n}{\boldsymbol{C}}$ be matrices of rank r such that $\boldsymbol{A} = \boldsymbol{BC}$. Then

$$\boldsymbol{BCBC} = \boldsymbol{AA} = k\boldsymbol{A} = k\boldsymbol{BC} = \boldsymbol{B}(k\underset{r\times r}{\boldsymbol{I}})\boldsymbol{C}.$$

Now $\boldsymbol{B}$ of full column rank implies $\boldsymbol{CBC} = k\underset{r\times r}{\boldsymbol{I}}\boldsymbol{C}$, and
$\boldsymbol{C}$ of full row rank implies $\boldsymbol{CB} = k\underset{r\times r}{\boldsymbol{I}}$.

Thus,

$$tr(\boldsymbol{A}) = tr(\boldsymbol{BC}) = tr(\boldsymbol{CB}) = tr(k\underset{r\times r}{\boldsymbol{I}}) = k \cdot tr(\underset{r\times r}{\boldsymbol{I}}) = k \cdot r = k \cdot rank(\boldsymbol{A}). \quad \square$$

Prove that $tr(\boldsymbol{I} - \boldsymbol{P_X}) = n - rank(\boldsymbol{X})$.

Proof:

We know $\boldsymbol{I} - \boldsymbol{P_X}$ is idempotent. Thus, $tr(\boldsymbol{I} - \boldsymbol{P_X}) = rank(\boldsymbol{I} - \boldsymbol{P_X})$.

We know $\boldsymbol{I} - \boldsymbol{P_X}$ is the orthogonal projection matrix onto $\mathcal{C}(\boldsymbol{X})^{\perp} = \mathcal{N}(\boldsymbol{X}')$.

Thus, $\mathcal{C}(\boldsymbol{I} - \boldsymbol{P_X}) = \mathcal{N}(\boldsymbol{X}')$, which has dimension $n - rank(\boldsymbol{X})$.

Thus, $rank(\boldsymbol{I} - \boldsymbol{P_X}) = n - rank(\boldsymbol{X})$. $\square$

Alternate Proof:

Because $\boldsymbol{P_X}$ is idempotent, $tr(\boldsymbol{P_X}) = rank(\boldsymbol{P_X})$.

Now note that $rank(\boldsymbol{P_X}) = rank(\boldsymbol{X})$ because

$$rank(\boldsymbol{P_X}) = rank(\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-}\boldsymbol{X}') \leq rank(\boldsymbol{X}) = rank(\boldsymbol{P_X X}) \leq rank(\boldsymbol{P_X}).$$

(This also follows from $\mathcal{C}(\boldsymbol{X}) = \mathcal{C}(\boldsymbol{P_X})$.)

Thus, $tr(\boldsymbol{I} - \boldsymbol{P_X}) = tr(\boldsymbol{I}) - tr(\boldsymbol{P_X}) = n - tr(\boldsymbol{P_X}) = n - rank(\boldsymbol{X})$. $\square$