Subspaces in Machine Learning
Proofs for the Rank of a Product of Matrices

Background: A matrix is rank P if there are P linearly independent columns (or rows). Vectors $\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_P$ are linearly independent if and only if

$$\sum_{i=1}^{P} \boldsymbol{v}_i c_i = \boldsymbol{0}$$

implies $c_i = 0, i = 1, 2, \ldots, c_P$. For convenience we may write this condition in matrix vector form as $\boldsymbol{V}\boldsymbol{c} = \boldsymbol{0}$ if and only if $\boldsymbol{c} = \boldsymbol{0}$ where

$$\boldsymbol{V} = \left[\begin{array}{cccc} \boldsymbol{v}_1 & \boldsymbol{v}_2 & \ldots & \boldsymbol{v}_P \end{array} \right]$$

and $\boldsymbol{c} = \left[\begin{array}{cccc} c_1 & c_2 & \ldots & c_P \end{array} \right]^T$.

Assume that $\boldsymbol{R} = \boldsymbol{A}\boldsymbol{B}$ where $\boldsymbol{R}$ is N-by-K, $\boldsymbol{A}$ is N-by-M, and $\boldsymbol{B}$ is M-by-K with $M \leq N$ and $M \leq K$. Let $\boldsymbol{B} = \left[\begin{array}{cccc} \boldsymbol{b}_1 & \boldsymbol{b}_2 & \ldots & \boldsymbol{b}_K \end{array} \right]$ have columns $\boldsymbol{b}_i$ and $\boldsymbol{R} = \left[\begin{array}{cccc} \boldsymbol{r}_1 & \boldsymbol{r}_2 & \ldots & \boldsymbol{r}_K \end{array} \right]$ have columns $\boldsymbol{r}_i$.

1. **Rank of a general product.** $\text{rank}(\boldsymbol{A}\boldsymbol{B}) \leq \min\{\text{rank}(\boldsymbol{A}), \text{rank}(\boldsymbol{B})\}$

   The ith column of $\boldsymbol{R}, \boldsymbol{r}_i$, is $\boldsymbol{A}\boldsymbol{b}_i$. A linear combination of some set of columns of $\boldsymbol{R}$ may be expressed in terms of $\boldsymbol{A}$ as

   $$\sum_j \boldsymbol{r}_j d_j = \sum_j \boldsymbol{A}\boldsymbol{b}_j d_j = \boldsymbol{A}\boldsymbol{c}$$

   where $\boldsymbol{c} = \sum_j \boldsymbol{b}_j d_j$ and the sum over j is taken with respect to any subset of the columns. Note that $\boldsymbol{c} = \boldsymbol{0}$ for $d_j \neq 0$ when the number of terms in the sum exceeds $\text{rank}(\boldsymbol{B})$. Hence $\text{rank}(\boldsymbol{R}) \leq \text{rank}(\boldsymbol{B})$. Similarly, writing $\sum_j \boldsymbol{r}_j d_j = \boldsymbol{A}\boldsymbol{c}$ indicates that $\text{rank}(\boldsymbol{R})$ is limited by $\text{rank}(\boldsymbol{A})$.

2. **Rank of a product of rank M matrices.** If $\text{rank}(\boldsymbol{A}) = M$ and $\text{rank}(\boldsymbol{B}) = M$, then $\text{rank}(\boldsymbol{R}) = M$. The converse is also true.

   First suppose $\text{rank}(\boldsymbol{R}) = M$. By the previous result we have $M \leq \min\{\text{rank}(\boldsymbol{A}), \text{rank}(\boldsymbol{B})\}$, but $\text{rank}(\boldsymbol{A})$ and $\text{rank}(\boldsymbol{B})$ are at most rank M (their smallest dimension). Thus $\text{rank}(\boldsymbol{A}) = \text{rank}(\boldsymbol{B}) = M$.

   Now assume $\text{rank}(\boldsymbol{A}) = \text{rank}(\boldsymbol{B}) = M$. We have $\text{rank}(\boldsymbol{R}) = M$ if and only if there is a set of M columns with indices $j = j_1, j_2, \ldots j_M$ for which

   $$\sum_{j=j_1,j_2,\ldots j_M} \boldsymbol{r}_j d_j = \boldsymbol{0}$$

implies $d_j = 0, j = j_1, j_2, \ldots j_M$. We have

$$\sum_{j=j_1,j_2,\ldots j_M} \boldsymbol{r}_j d_j = \boldsymbol{A} \sum_j \boldsymbol{b}_j d_j = \boldsymbol{A}\boldsymbol{c}$$

where $\boldsymbol{c} = \sum_j \boldsymbol{b}_j d_j$. Since $\text{rank}(\boldsymbol{A}) = M$, we know that $\boldsymbol{A}\boldsymbol{c} = \boldsymbol{0}$ if and only if $\boldsymbol{c} = \boldsymbol{0}$. Now note that

$$\boldsymbol{c} = \begin{bmatrix} \boldsymbol{b}_{j_1} & \boldsymbol{b}_{j_2} & \cdots & \boldsymbol{b}_{j_M} \end{bmatrix} \boldsymbol{d} = \tilde{\boldsymbol{B}}\boldsymbol{d}$$

Since $\text{rank}(\boldsymbol{B}) = M$, there is a set of M columns $j = j_1, j_2, \ldots j_M$ such that $\boldsymbol{c} = \boldsymbol{0}$ if and only if $\boldsymbol{d} = \boldsymbol{0}$, which proves that $\text{rank}(\boldsymbol{R}) = M$.