Solving the Least-Squares Problem Using Geometry Proofs: Invertibility of Full Rank and Positive Definite Matrices

©Barry Van Veen 2019

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

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

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

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

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

The columns of a rank P, N-by-P matrix span a P-dimensional subspace of $\mathbb{R}^N$.

1. **$\boldsymbol{A}^T\boldsymbol{A}$ is full rank.** If $\boldsymbol{A}$ is N-by-P with $P \leq N$ and rank$\{\boldsymbol{A}\} = P$, then $\boldsymbol{B} = \boldsymbol{A}^T\boldsymbol{A}$ (P-by-P) has rank P.

   We will prove this result by contradiction. Suppose $\boldsymbol{B}$ has rank less than P. This implies there is a vector $\boldsymbol{v} \neq \mathbf{0}$ such that $\boldsymbol{B}\boldsymbol{v} = \mathbf{0}$. Since $\boldsymbol{B}\boldsymbol{v} = \mathbf{0}$, we have $\boldsymbol{v}^T\boldsymbol{B}\boldsymbol{v} = 0$ for some $\boldsymbol{v} \neq \mathbf{0}$. Note that $\boldsymbol{v}^T\boldsymbol{B}\boldsymbol{v} = \boldsymbol{v}^T\boldsymbol{A}^T\boldsymbol{A}\boldsymbol{v} = (\boldsymbol{A}\boldsymbol{v})^T\boldsymbol{A}\boldsymbol{v} = \boldsymbol{y}^T\boldsymbol{y}$ where we define $\boldsymbol{y} = \boldsymbol{A}\boldsymbol{v}$. But rank$\{\boldsymbol{A}\} = P$, so there is no $\boldsymbol{v} \neq \mathbf{0}$ for which $\boldsymbol{A}\boldsymbol{v} = \mathbf{0}$. Thus, if $\boldsymbol{v} \neq \mathbf{0}$, then $\boldsymbol{y} \neq \mathbf{0}$. If $\boldsymbol{y} \neq \mathbf{0}$, then $\boldsymbol{y}^T\boldsymbol{y} = \boldsymbol{v}^T\boldsymbol{B}\boldsymbol{v} > 0$. Hence there is a contradiction and rank$\{\boldsymbol{B}\} = P$.

2. **Positive Definite Matrices are Full Rank.** If a P-by$P$ matrix $\boldsymbol{Q}$ is positive definite, that is, $\boldsymbol{v}^T\boldsymbol{Q}\boldsymbol{v} > 0$ for all $\boldsymbol{v} \neq \mathbf{0}$, then rank$\{\boldsymbol{Q}\} = P$.

   The proof is similar to the previous one. Suppose rank$\{\boldsymbol{Q}\} < P$. This implies there is a vector $\boldsymbol{v} \neq \mathbf{0}$ such that $\boldsymbol{Q}\boldsymbol{v} = \mathbf{0}$. If $\boldsymbol{Q}\boldsymbol{v} = \mathbf{0}$, then $\boldsymbol{v}^T\boldsymbol{Q}\boldsymbol{v} = \boldsymbol{v}^T\mathbf{0} = 0$, which contradicts the assumption that $\boldsymbol{Q}$ is positive definite. Thus rank$\{\boldsymbol{Q}\} = P$.

3. **Full Rank Square Matrices Are Invertible.** If $\boldsymbol{B}$ is P-by-P and rank$\{\boldsymbol{B}\} = P$, then $\boldsymbol{B}$ is invertible. That is, there exists $\boldsymbol{B}^{-1}$ satisfying $\boldsymbol{B}\boldsymbol{B}^{-1} = \boldsymbol{I}$.

   Since rank$\{\boldsymbol{B}\} = P$, the columns of $\boldsymbol{B}$ span $\mathbb{R}^P$ and any vector $\boldsymbol{z} \in \mathbb{R}^P$ can be written as a weighted combination of the columns of $\boldsymbol{B}$, that is, $\boldsymbol{z} = \boldsymbol{B}\boldsymbol{v}$.

   Now, let $\boldsymbol{e}_i \in \mathbb{R}^P$ be the vector of all zeros except for a one in the i^{th} row, e.g., $\boldsymbol{e}_2 = \left[\begin{array}{ccccc} 0 & 1 & 0 & \cdots & 0 \end{array} \right]^T$. Since the columns of $\boldsymbol{B}$ span $\mathbb{R}^P$, there is a vector $\boldsymbol{v}_i$ so that $\boldsymbol{e}_i = \boldsymbol{B}\boldsymbol{v}_i, i = 1, 2, \ldots, P$. Concatenating these relationships for $i = 1, 2, \ldots, P$ gives

$$\begin{bmatrix} \boldsymbol{e}_1 & \boldsymbol{e}_2 & \cdots & \boldsymbol{e}_P \end{bmatrix} = \boldsymbol{B} \begin{bmatrix} \boldsymbol{v}_1 & \boldsymbol{v}_2 & \cdots & \boldsymbol{v}_P \end{bmatrix}$$

But $\begin{bmatrix} \boldsymbol{e}_1 & \boldsymbol{e}_2 & \cdots & \boldsymbol{e}_P \end{bmatrix} = \boldsymbol{I}$, so thus by definition $\begin{bmatrix} \boldsymbol{v}_1 & \boldsymbol{v}_2 & \cdots & \boldsymbol{v}_P \end{bmatrix} = \boldsymbol{B}^{-1}$.