Given the Krylov subspace

$$\mathcal{K}(A, q_1, n) := \text{span}\{q_1, Aq_1, \dots, A^{n-1}q_1\} \tag{1}$$

and the Lanczos Algorithm

$$q_j = r_{j-1}/\beta_{j-1}, \tag{2}$$

$$\alpha_j = q_j^t A q_j \tag{3}$$

$$r_j = (A - \alpha_j I)q_j - \beta_{j-1}q_{j-1} \tag{4}$$

$$\beta_j = ||r_j||_2 \tag{5}$$

for $j = 1, \dots, M$ with $r_0 = q_1$, $\beta_0 = 1$, $q_0 = 0$, and M is the smallest positive integer such that $\beta_M = 0$.

$\boxed{\textbf{Theorem}}$: Let $A \in \mathbb{R}^{n\times n}$ be symmetric and assume $q_1 \in \mathbb{R}^n$ with $||q_1||_2 = 1$. Then the Lanczos iteration runs until $j = m$ where $m = \text{rank}(\{\mathcal{K}(A, q_1, n)\})$. Moreover, for $j = 1, \dots, m$ we have

$$AQ_j = Q_j T_j + r_j e_j^t \tag{6}$$

where

$$T_j = \begin{bmatrix} \alpha_1 & \beta_1 & \cdots & 0 & 0 \\ \beta_1 & \alpha_2 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & \alpha_{j-1} & \beta_{j-1} \\ 0 & 0 & \cdots & \beta_{j-1} & \alpha_j \end{bmatrix} \in \mathbb{R}^{j\times j} \tag{7}$$

$e_j = [0, 0, \dots, 1]^t \in \mathbb{R}^j$ and $Q_j = [q_1|\cdots|q_j]$ has orthonormal columns that span $\mathcal{K}(A, q_1, j)$.

$\boxed{\textbf{Proof}}$:

(1) We prove the statements by using the mathematical induction on j.

For $j = 1$, by (4) we have $Aq_1 = r_1 + \beta_0 q_0 + \alpha_1 q_1 = \alpha_1 q_1 + r_1 e_1^t$, then (6) followed by $Q_1 = q_1$ and $T_1 = [\alpha_1]$. The Q_2 has orthonormal columns is due to $q_1^t q_1 = ||q_1||_2^2 = 1$, $q_2^t q_2 = ||q_2||_2^2 = ||r_1/\beta_1||_2^2 = 1$, and $q_1^t q_2 = [q_1^t(A - \alpha_1 I)q_1 - \beta_0 q_1^t q_0]/\beta_1 = [q_1^t A q_1 - \alpha_1 q_1^t q_1]/\beta_1 = 0$. And by (2) and (4) we have $Aq_1 = \alpha_1 q_1 + r_1 = \alpha_1 q_1 + \beta_1 q_2$, so $\text{span}\{q_1, q_2\} = \text{span}\{q_1, Aq_1\} = \mathcal{K}(A, q_1, 2)$.

Now assume that it holds for $j \le k$, where $k \le M - 1$ (i.e., $\beta_k \neq 0$).

We have $AQ_k = Q_k T_k + r_j e_k^t$, $Q_k^t Q_k = I_k$, and $\text{range}(Q_j) = \mathcal{K}(A, q_1, j)$ for $j \le k$. When $j = k + 1$, we have

$$\begin{aligned} Q_{k+1}T_{k+1} + r_{k+1}e_{k+1}^t &= \begin{bmatrix} Q_k & q_{k+1} \end{bmatrix} \begin{bmatrix} T_k & \beta_k e_k \\ \beta_k e_k^t & \alpha_{k+1} \end{bmatrix} + r_{k+1} \begin{bmatrix} 0_{k\times 1} & 1 \end{bmatrix} \\ &= \begin{bmatrix} Q_k T_k + \beta_k q_{k+1} e_k^t & \beta_k Q_k e_k + \alpha_{k+1} q_{k+1} + r_{k+1} \end{bmatrix} \\ &= \begin{bmatrix} Q_k T_k + r_k e_k^t & \beta_k q_k + \alpha_{k+1} q_{k+1} + r_{k+1} \end{bmatrix} \\ &= \begin{bmatrix} AQ_k & Aq_{k+1} \end{bmatrix} \\ &= AQ_{k+1} \end{aligned}$$

where the third and forth equalities are due to (2), (4), and induction hypothesis, so we had proven (6) for $j = k + 1$.

To prove $Q_{k+1}^t Q_{k+1} = I_{k+1}$, it suffices to prove $q_i^t q_{k+1} = 0$ for $i \le k$ and $q_{k+1}^t q_{k+1} = 1$. The latter one is obvious (by (2)) and the former one is more complicated. For $i = k$, it is clear that

$$\begin{aligned} q_k^t q_{k+1} &= (q_k^t A q_k - \alpha_k q_k^t q_k - \beta_{k-1} q_k^t q_{k-1})/\beta_k \\ &= (q_k^t A q_k - \alpha_k)/\beta_k \\ &= 0. \end{aligned}$$

For $i = k-1$, we have

$$\begin{aligned} q_{k-1}^t q_{k+1} &= (q_{k-1}^t A q_k - \alpha_k q_{k-1}^t q_k - \beta_{k-1} q_{k-1}^t q_{k-1})/\beta_k \\ &= (q_{k-1}^t A q_k - \beta_{k-1})/\beta_k \\ &= 0, \end{aligned}$$

where

$$\begin{aligned} q_{k-1}^t A q_k &= (A q_{k-1})^t q_k \\ &= (r_{k-1} + \alpha_{k-1} q_{k-1} + \beta_{k-1} q_{k-2})^t q_k \\ &= r_{k-1}^t q_k = r_{k-1}^t r_{k-1}/\beta_{k-1} \\ &= \beta_{k-1}. \end{aligned}$$

And for $i \le k-2$ we have $q_i^t q_{k+1} = (q_i^t A q_k - \alpha_k q_i^t q_k - \beta_{k-1} q_i^t q_{k-1})/\beta_k = ((Aq_i)^t) q_k/\beta_k = 0$, where the last equality we used $q_i \in \mathcal{K}(A, q_1, k-2) = \operatorname{span}\{q_1, Aq_1, \ldots, A^{k-3} q_1\}$ so $Aq_i \in \operatorname{span}\{Aq_1, A^2 q_1, \ldots, A^{k-2} q_1\} \subset \mathcal{K}(A, q_1, k-1)$, i.e., Aq_i is a linear combination of $\{q_1, q_2, \ldots, q_{k-1}\}$, so $(Aq_i)^t q_k = 0$.

It is clearly that $\operatorname{rank}(\mathcal{K}(A, q_1, k+1)) \le k+1$. Since $q_1, \ldots, q_{k+1}$ are linearly independent, $\operatorname{span}\{q_1, \ldots, q_k\} \subset \mathcal{K}(A, q_1, k) \subset \mathcal{K}(A, q_1, k+1)$, and $q_{k+1} = (Aq_k - \alpha_k q_k - \beta_{k-1} q_{k-1})/\beta_k \in \operatorname{span}\{Aq_k, q_k, q_{k-1}\} \subset \operatorname{span}\{q_1, Aq_1, \ldots, A^k q_1\} = \mathcal{K}(A, q_1, k+1)$, so $\operatorname{span}\{q_1, \ldots, q_{k+1}\} = \mathcal{K}(A, q_1, k+1)$.

(2) We show that $M = m$, in fact, for $j = M-1$, the above results tell us $\operatorname{span}\{q_1, \ldots, q_M\} = \mathcal{K}(A, q_1, M) \subset \mathcal{K}(A, q_1, n)$, thus $m \ge M$.

Now, since M is finite, so for $j = M$ we have $AQ_M = Q_M T_M$, thus, Aq_M is a linear combination of $\{q_1, \ldots, q_M\}$ for $i \le M$. But

$$\begin{aligned} \mathcal{K}(A, q_1, M+1) &= \operatorname{span}\{q_1, Aq_1, \ldots, A^M q_1\} \\ &= \operatorname{span}\{q_1, A\mathcal{K}(A, q_1, M)\} \\ &= \operatorname{span}\{q_1, Aq_1, \ldots, Aq_M\} \\ &= \operatorname{span}\{q_1, Aq_1, \ldots, Aq_{M-1}\} \\ &= \operatorname{span}\{q_1, A\mathcal{K}(A, q_1, M-1)\} \\ &= \operatorname{span}\{q_1, Aq_1, \ldots, A^{M-1} q_1\} \\ &= \mathcal{K}(A, q_1, M) \end{aligned}$$

and use the induction sense we can show that

$$\begin{aligned} \mathcal{K}(A, q_1, i+1) &= \operatorname{span}\{q_1, Aq_1, \ldots, A^i q_1\} \\ &= \operatorname{span}\{q_1, A\mathcal{K}(A, q_1, i)\} \\ &= \operatorname{span}\{q_1, A\mathcal{K}(A, q_1, i-1)\} \\ &= \operatorname{span}\{q_1, Aq_1, \ldots, A^{i-1} q_1\} \\ &= \mathcal{K}(A, q_1, i) \end{aligned}$$

for $i \ge M+1$ provided $\mathcal{K}(A, q_1, i) = \mathcal{K}(A, q_1, i-1)$. So $\mathcal{K}(A, q_1, M) = \mathcal{K}(A, q_1, n)$ has rank m, i.e., $M \ge m$.