Theory of Matrices

Chapter 5. Matrix Transformations and Factorizations

5.7. *QR* Factorization—Proofs of Theorems

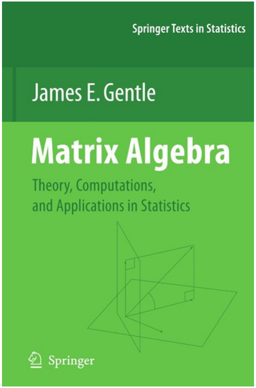

Table of contents

Theorem 5.7.A

Theorem 5.7.A. Let $\{a_1, a_2, \ldots, a_k\}$ be a linearly independent set of vectors. There exists unique scalars x_{ij} where $1 \le j \le k$ and $0 < i < j$ such that the k vectors

$$\begin{aligned} b_1 &= a_1 \\ b_2 &= a_2 - x_{12} b_1 \\ &\vdots \\ b_j &= a_j - x_{j-1,j} b_{j-1} - x_{j-2,j} b_{j-2} - \cdots - x_{1j} b_1 \\ &\vdots \\ b_m &= a_m - x_{m-1,m} b_{m-1} - x_{m-2,m} b_{m-2} - \cdots - x_{1m} b_1 \end{aligned}$$

form an orthogonal set. The vectors $b_1, b_2, \ldots, b_k$ are nonzero and $x_{ij} = \langle a_j, b_i \rangle / \langle b_i, b_i \rangle$ for $1 \le j \le m$ and $0 < i < m$.

Theorem 5.7.A (continued 1)

Proof (continued). We prove this by induction on k. The result is trivial for $k = 1$ since we have $b_1 = a_1$ and there are no x_{ij}. Suppose the result holds for $m = k - 1$. That is, suppose there exists unique x_{ij} for $1 \leq j \leq k - 1$ and $0 < i < j$ such that $b_1, b_2, \ldots, b_{k-1}$ defined as given in the statement of the theorem form an orthogonal set of nonzero vectors where $x_{ij} = \langle a_j, b_i \rangle / \langle b_i, b_i \rangle$. We define b_k as given in the statement of the theorem and we must show that b_k is nonzero and that the coefficients x_{ij} are as claimed for $j = k$ and $0 < i < j$. Notice that for $0 < i < j$ we have

$$\langle b_k, b_i \rangle = \langle a_k - x_{k-1,\,k} b_{k-1} - x_{k-2,\,k} b_{k-2} - \cdots - x_{1k} b_1, b_i \rangle$$

$$= \langle a_k, b_i \rangle - x_{k-1,\,k} \langle b_{k-1}, b_i \rangle - \cdots - x_{i,k} \langle b_i, b_i \rangle - \cdots - x_{1k} \langle b_1, b_i \rangle$$

$$= \langle a_k, b_i \rangle - x_{ik} \langle b_i, b_i \rangle.$$

Theorem 5.7.A (continued 1)

Proof (continued). We prove this by induction on k. The result is trivial for $k = 1$ since we have $b_1 = a_1$ and there are no x_{ij}. Suppose the result holds for $m = k - 1$. That is, suppose there exists unique x_{ij} for $1 \le j \le k-1$ and $0 < i < j$ such that $b_1, b_2, \ldots, b_{k-1}$ defined as given in the statement of the theorem form an orthogonal set of nonzero vectors where $x_{ij} = \langle a_j, b_i \rangle / \langle b_i, b_i \rangle$. We define b_k as given in the statement of the theorem and we must show that b_k is nonzero and that the coefficients x_{ij} are as claimed for $j = k$ and $0 < i < j$. Notice that for $0 < i < j$ we have

$$\langle b_k, b_i \rangle = \langle a_k - x_{k-1,k} b_{k-1} - x_{k-2,k} b_{k-2} - \cdots - x_{1k} b_1, b_i \rangle$$

$$= \langle a_k, b_i \rangle - x_{k-1,k} \langle b_{k-1}, b_i \rangle - \cdots - x_{i,k} \langle b_i, b_i \rangle - \cdots - x_{1k} \langle b_1, b_i \rangle$$

$$= \langle a_k, b_i \rangle - x_{ik} \langle b_i, b_i \rangle.$$

So $\langle b_k, b_i \rangle = 0$ if and only if $\langle a_k, b_i \rangle - x_{ik} \langle b_i, b_i \rangle = 0$ or, since $b_i \ne 0$ by the induction hypothesis, $\langle b_k, b_i \rangle = 0$ (that is, $b_1, b_2, \ldots, b_{k-1}, b_k$ form an orthogonal set) if and only if $x_{ik} = \langle a_k, b_i \rangle / \langle b_i, b_i \rangle$.

Theorem 5.7.A (continued 1)

Proof (continued). We prove this by induction on k. The result is trivial for $k = 1$ since we have $b_1 = a_1$ and there are no x_{ij}. Suppose the result holds for $m = k - 1$. That is, suppose there exists unique x_{ij} for $1 \le j \le k - 1$ and $0 < i < j$ such that $b_1, b_2, \ldots, b_{k-1}$ defined as given in the statement of the theorem form an orthogonal set of nonzero vectors where $x_{ij} = \langle a_j, b_i\rangle / \langle b_i, b_i\rangle$. We define b_k as given in the statement of the theorem and we must show that b_k is nonzero and that the coefficients x_{ij} are as claimed for $j = k$ and $0 < i < j$. Notice that for $0 < i < j$ we have

$$\langle b_k, b_i\rangle = \langle a_k - x_{k-1,\,k} b_{k-1} - x_{k-2,\,k} b_{k-2} - \cdots - x_{1k} b_1, b_i\rangle$$

$$= \langle a_k, b_i\rangle - x_{k-1,\,k}\langle b_{k-1}, b_i\rangle - \cdots - x_{i,k}\langle b_i, b_i\rangle - \cdots - x_{1k}\langle b_1, b_i\rangle$$

$$= \langle a_k, b_i\rangle - x_{ik}\langle b_i, b_i\rangle.$$

So $\langle b_k, b_i\rangle = 0$ if and only if $\langle a_k, b_i\rangle - x_{ik}\langle b_i, b_i\rangle = 0$ or, since $b_i \ne 0$ by the induction hypothesis, $\langle b_k, b_i\rangle = 0$ (that is, $b_1, b_2, \ldots, b_{k-1}, b_k$ form an orthogonal set) if and only if $x_{ik} = \langle a_k, b_i\rangle / \langle b_i, b_i\rangle$.

Theorem 5.7.A (continued 2)

Proof (continued). So the scalars x_{ij} for $j = k$ and $0 < i < j$ must satisfy $x_{ik} = \langle a_k, b_i\rangle / \langle b_i, b_i\rangle$. So by induction, the x_{ij} are as claimed for $1 \le j \le k$ and $0 < i < j$.

We just need to show that b_k is nonzero. From the definition of $b_1, b_2, \ldots, b_k$ we can express:

- b_1 in terms of a_1 with a coefficient of 1 for a_1,
- b_2 in terms of a_2 and a_1 with a coefficient of 1 for a_2,
- b_3 in terms of a_3, a_2, and a_1 with a coefficient of 1 for a_3, $\ldots$, and
- b_k in terms of $a_k, a_{k-1}, \ldots, a_1$ with a coefficient of 1 for a_k.

Theorem 5.7.A (continued 2)

Proof (continued). So the scalars x_{ij} for $j = k$ and $0 < i < j$ must satisfy $x_{ik} = \langle a_k, b_i \rangle / \langle b_i, b_i \rangle$. So by induction, the x_{ij} are as claimed for $1 \le j \le k$ and $0 < i < j$.

We just need to show that b_k is nonzero. From the definition of $b_1, b_2, \ldots, b_k$ we can express:

- b_1 in terms of a_1 with a coefficient of 1 for a_1,
- b_2 in terms of a_2 and a_1 with a coefficient of 1 for a_2,
- b_3 in terms of a_3, a_2, and a_1 with a coefficient of 1 for a_3, $\ldots$, and
- b_k in terms of $a_k, a_{k-1}, \ldots, a_1$ with a coefficient of 1 for a_k.

Since $\{a_1, a_2, \ldots, a_k\}$ is a linearly independent set and when expressing b_k as a linear combination of these vectors the coefficient of a_k is nonzero then b_k is nonzero. So the result holds for $m = k$ and therefore holds in general by induction. □

Theorem 5.7.A (continued 2)

Proof (continued). So the scalars x_{ij} for $j = k$ and $0 < i < j$ must satisfy $x_{ik} = \langle a_k, b_i \rangle / \langle b_i, b_i \rangle$. So by induction, the x_{ij} are as claimed for $1 \leq j \leq k$ and $0 < i < j$.

We just need to show that b_k is nonzero. From the definition of $b_1, b_2, \ldots, b_k$ we can express:

- b_1 in terms of a_1 with a coefficient of 1 for a_1,
- b_2 in terms of a_2 and a_1 with a coefficient of 1 for a_2,
- b_3 in terms of a_3, a_2, and a_1 with a coefficient of 1 for a_3, $\ldots$, and
- b_k in terms of $a_k, a_{k-1}, \ldots, a_1$ with a coefficient of 1 for a_k.

Since $\{a_1, a_2, \ldots, a_k\}$ is a linearly independent set and when expressing b_k as a linear combination of these vectors the coefficient of a_k is nonzero then b_k is nonzero. So the result holds for $m = k$ and therefore holds in general by induction. □

Theorem 5.7.B

Theorem 5.7.B. Let A be an $n \times m$ matrix of full column rank (that is, $\text{rank}(A) = m$). Then there is $n \times m$ matrix B, where the columns of B are mutually orthogonal and nonzero, and $m \times m$ upper triangular matrix X, with all diagonal entries 1, such that $A = BX$.

Proof. Denote the (linearly independent) columns of A as $a_1, a_2, \ldots, a_m$. Recursively define (column) vectors $b_1, b_2, \ldots, b_m$ as

$$\begin{aligned} b_1 &= a_1 \\ b_2 &= a_2 - x_{12} b_1 \\ b_3 &= a_3 - x_{23} b_2 - x_{13} b_1 \\ &\vdots \\ b_j &= a_j - x_{j-1,j} b_{j-1} - x_{j-2,j} b_{j-2} - \cdots - x_{1j} b_1 \\ &\vdots \\ b_m &= a_m - x_{m-1,m} b_{m-1} - x_{m-2,m} b_{m-2} - \cdots - x_{1m} b_1 \end{aligned}$$

Theorem 5.7.B

Theorem 5.7.B. Let A be an $n \times m$ matrix of full column rank (that is, $\text{rank}(A) = m$). Then there is $n \times m$ matrix B, where the columns of B are mutually orthogonal and nonzero, and $m \times m$ upper triangular matrix X, with all diagonal entries 1, such that $A = BX$.

Proof. Denote the (linearly independent) columns of A as $a_1, a_2, \ldots, a_m$. Recursively define (column) vectors $b_1, b_2, \ldots, b_m$ as

$$\begin{aligned}
b_1 &= a_1 \\
b_2 &= a_2 - x_{12} b_1 \\
b_3 &= a_3 - x_{23} b_2 - x_{13} b_1 \\
&\vdots \\
b_j &= a_j - x_{j-1,j} b_{j-1} - x_{j-2,j} b_{j-2} - \cdots - x_{1j} b_1 \\
&\vdots \\
b_m &= a_m - x_{m-1,m} b_{m-1} - x_{m-2,m} b_{m-2} - \cdots - x_{1m} b_1
\end{aligned}$$

Theorem 5.7.B (continued 1)

Proof (continued). ... where $x_{ij} = \langle a_j, b_i \rangle / \langle b_i, b_i \rangle$. The $b_1, b_2, \ldots, b_m$ are mutually orthogonal and each is nonzero by Theorem 5.7.A. We can rearrange the equations defining the b_i's to express the a_j's as linear combinations of the b_i's:

$$\begin{aligned} a_1 &= b_1 \\ a_2 &= b_2 + x_{12} b_1 \\ &\vdots \\ a_j &= b_j + x_{j-1,j} b_{j-1} + x_{j-2,j} b_{j-2} + \cdots + x_{1j} b_1 \\ &\vdots \\ a_m &= b_m + x_{m-1,m} b_{m-1} + x_{m-2,m} b_{m-2} + \cdots + x_{1m} b_1 \qquad (*) \end{aligned}$$

where $x_{ij} = \langle a_j, b_i \rangle / \langle b_i, b_i \rangle$.

Theorem 5.7.B (continued 2)

Theorem 5.7.B. Let A be an $n \times m$ matrix of full column rank (that is, rank$(A) = m$). Then there is $n \times m$ matrix B, where the columns of B are mutually orthogonal and nonzero, and $m \times m$ upper triangular matrix X, with all diagonal entries 1, such that $A = BX$.

Proof (continued). Define $n \times m$ matrix B with columns $b_1, b_2, \ldots, b_m$ and $m \times m$ matrix X with entry (i, j) as x_{ij} for $i < j$, $x_{ii} = 1$, and $x_{ij} = 0$ for $i > j$. So the columns of B are orthogonal and X is upper triangular with diagonal entries of 1. From $(*)$, we see that $A = BX$, as claimed. □