Suppose $A$ is an $m \times n$ matrix of approximate rank $k$, and that we have identified two index sets $I_s$ and $J_s$ such that the matrices

$$C = A(:, J_s)$$
$$R = A(I_s, :)$$

hold $k$ columns/rows that approximately span the column/row space of $A$. You may assume that $C$ and $R$ both have rank $k$ (in other words, the index vectors $J_s$ and $I_s$ are not very bad). Then

$$A \approx CC^\dagger AR^\dagger R,$$

and the optimal choice for the “$U$” factor in the CUR decomposition is,

$$U = C^\dagger AR^\dagger.$$

Set $X = CC^\dagger$.

(a) Suppose that $C$ has SVD

$$C = UDV^*.$$

Prove that $X = UU^*$.

**Solution:** Let $C = UDV^*$. Then

$$X = CC^\dagger$$
$$= UDV^* (UDV^*)^\dagger$$
$$= UDV^* VD^\dagger U^*$$
$$= UDD^\dagger U^*$$
$$= UDD^{-1} U^*$$
$$= UU^*.$$

Note that $D^\dagger = D^{-1}$ since $C$ is $m \times k$ and of rank $k$.

(b) Suppose that $C$ has the QR factorization

$$CP = QS$$

Prove that $X = QQ^*$. (Observe that $S$ is necessarily invertible, since $C$ has rank $k$. You can then prove that $C^\dagger = PS^{-1}Q^*$.)
Solution: \( CP = QS \implies C = QSP^* \) since \( PP^* = P^*P = I \). Then
\[
X = CC^\dagger = QSP^*(QSP^*)^\dagger = QSP^*PS^\dagger Q^* = QSS^\dagger Q^* = QSS^{-1}Q^* = QQ^*
\]

Note that \( S^\dagger = S^{-1} \) since \( C \) is \( m \times k \) and of rank \( k \).

(c) Prove that \( X \) is the orthogonal projection onto \( \text{Col}(C) \).

Solution: First, in order for \( X \) to be an orthogonal projection, it must satisfy \( X = X^* \) and \( X^2 = X \).

Let \( C = UDV^* \) be the SVD of \( C \) as in part(a). Then \( X = CC^\dagger = UU^* \) and \( X^* = (UU^*)^* = UU^* = X \).

Moreover, \( XX^* = X^2 = (UU^*)(UU^*) = UU^* = X \), and so \( X \) is an orthogonal projection. It is also straightforward to check \( ||X|| = 1 \) since \( U \) is orthogonal.

Now, it is left to show that \( X \) projects onto \( \text{Col}(C) \). Recall the definition of the Moore-Penrose pseudo-inverse: \( C^\dagger = (C^*C)^{-1}C^* \), where \( C \) is \( m \times k \) with \( k \) linearly independent columns and decompose the space \( C = \text{ran}(C) \oplus \ker(C^*) \). Let \( v \in \text{ran}(C) = \text{col}(C) \), then there exists a \( u \) such that \( v = Cu \). Furthermore,
\[
Xv = CC^\dagger v = C(C^*C)^{-1}C^*v = C(C^*C)^{-1}C^*Cu = Cu = v.
\]

Suppose \( w \in \ker(C^*) \), then \( C^*w = 0 \).
\[
Xx = CC^\dagger w = C(C^*C)^{-1}C^*w = 0.
\]
Since \( X \) projects element from the range of \( C \) to itself and elements from the kernel to the 0 element, \( X \) is a projection operator onto \( \text{col}(C) \).

(d) Suppose that \( A \) has precisely rank \( k \) and that \( C \) and \( R \) are both of rank \( k \). Prove that then
\[
C^\dagger AR^\dagger = (A(I_s, J_s))^{-1}.
\]

Solution: Let \( \text{rank}(A) = \text{rank}(C) = \text{rank}(R) = k \) and recall that \( C = A(:, J_s) \) and \( R = A(I_s, :) \). Thus, \( C(I_s, :) = A(I_s, J_s) \) is a \( k \times k \) matrix of rank \( k \), implying \( A(I_s, J_s) \) is invertible. Moreover, since \( \text{rank}(A) = k \), we have from class that
\[ A = CA(I_s, J_s)^{-1}R, \]

the double sided ID.

We will digress for a moment and reprove it here. Since \( A \) is precisely of rank \( k \), it admits a factorization

\[ A = CZ, \]

where \( C = A(:, J_s) \) and \( Z \) contains some the \( k \times k \) identity matrix as a sub-matrix as well as the expansion coefficients used to build \( A \) from the skeleton columns contained in \( C \). \( A \) also admits the factorization

\[ A = XR \]

where \( R = A(I_s, :) \) consisting of \( k \) rows of \( A \), where \( X \) also contains the \( k \times k \) identity with a different set of expansion coefficients used to build \( A \). Taking the \( I_s \) rows of the Column-ID, we have

\[ A(I_s, :) = C(I_s, :)Z = A(I_s, J_s)Z, \]

it must be the case that

\[ Z = (A(I_s, J_s))^{-1}A(I_s, :). \]

Thus,

\[ A = CZ = C(A(I_s, J_s))^{-1}A(I_s, :) = C(A(I_s, J_s))^{-1}R. \]

Now, left multiplying both sides by \( C^\dagger \) and right multiplying by \( R^\dagger \) yields

\[ C^\dagger AR^\dagger = C^\dagger CA(I_s, J_s)^{-1}RR^\dagger = A(I_s, J_s)^{-1} \]

since \( C^\dagger \) is the left inverse of \( C \) and \( R^\dagger \) is the right inverse of \( R \).