Interpolative Decomposition

1. Interpolative Decomposition Continued

Theorem 1. *Let* $\mathbf{A}$ *be an* $m \times n$ *matrix of exact rank k. The* $\mathbf{A}$ *admits a factorization*

$$\underset{m \times n}{\mathbf{A}} = \underset{m \times k}{\mathbf{A}(:, \mathbf{J}_s)} \; \underset{k \times n}{\mathbf{Z}}. \tag{1}$$

Where $\mathbf{J}_s$ *is an index vector of length k contained in [1, 2, ..., n], and where* $\mathbf{Z}(:, \mathbf{J}_s) = \mathbf{I}_k$ *and* $max_{i,j} \|\mathbf{Z}(i,j)\| \leq 1$

1.1. **Proof.** Case 1: Suppose m = k so $\mathbf{A}$ is k $\times$ n. Pick $\mathbf{J}_s$ so that $\|det(\mathbf{A}(:, \mathbf{J}_s))\|$ is maximized. Let $\mathbf{J}_r$ denote the remaining indices, so

$$\mathbf{J}_s \cup \mathbf{J}_r = \{1, 2, 3, \dots, n\} \tag{2}$$

Where $\cup$ is the disjoint union. In other words,

$$[\mathbf{J}_s \quad \mathbf{J}_r] = [\mathbf{j}_1 \quad \mathbf{j}_2 \quad \cdots \quad \mathbf{j}_k \quad \mathbf{j}_{k+1} \quad \cdots \quad \mathbf{j}_n] \tag{3}$$

$$\mathbf{AP} = \mathbf{A}(:, \mathbf{J}) = [\mathbf{A}(:, \mathbf{J}_s) \quad \mathbf{A}(:, \mathbf{J}_r)] \tag{4}$$

For some permutation matrix $\mathbf{P}$

$$\mathbf{A} = \mathbf{A}(:, \mathbf{J})\,[\mathbf{I}_k \qquad \mathbf{A}(:, \mathbf{J}_s)^{-1}\mathbf{A}(:, \mathbf{J}_r)]\mathbf{P}^* \tag{5}$$

With the following

$$\mathbf{C} = \mathbf{A}(:, \mathbf{J})$$
$$\mathbf{Z} = [\mathbf{I}_k \qquad \mathbf{A}(:, \mathbf{J}_s)^{-1}\mathbf{A}(:, \mathbf{J}_r)]\mathbf{P}^*$$
$$\mathbf{T} := \mathbf{A}(:, \mathbf{J}_s)^{-1}\mathbf{A}(:, \mathbf{J}_r)$$

It remains to be shown that $\|\mathbf{T}(i,j)\| \leq 1$ Recall Cramer's Rule which states that given a k $\times$ k matrix $\mathbf{B}$ and the equation $\mathbf{Bx} = \mathbf{y}$, that $\mathbf{x}$ can be found by the following method.

$$\mathbf{x}(i) = \frac{det[\mathbf{b}_1 \quad \mathbf{b}_2 \cdots \mathbf{b}_{i-1} \quad \mathbf{y} \quad \mathbf{b}_{i+1} \cdots \mathbf{b}_k]}{det[\mathbf{b}_1 \quad \mathbf{b}_2 \quad \cdots \quad \mathbf{b}_k]}$$

where $\mathbf{b}_j$ is the j^{th} column of $\mathbf{B}$.

$$\mathbf{A} = \begin{bmatrix} | & | & & | \\ \mathbf{C}_1 & \mathbf{C}_2 & \cdots & \mathbf{C}_n \\ | & | & & | \end{bmatrix} \tag{6}$$

Then $\mathbf{A}(:, \mathbf{J}_s)\,\mathbf{T} = \mathbf{A}(:, \mathbf{J}_r)$. $\mathbf{T}$ is defined as the solution to this equation.

$$\begin{bmatrix} | & | & & | \\ \mathbf{C}_{j1} & \mathbf{C}_{j2} & \cdots & \mathbf{C}_{jk} \\ | & | & & | \end{bmatrix} \mathbf{T} = \begin{bmatrix} | & | & & | \\ \mathbf{C}_{jk+1} & \mathbf{C}_{jk+2} & \cdots & \mathbf{C}_{jn} \\ | & | & & | \end{bmatrix} \tag{7}$$

Cramer's Rule provides formulas for each $\mathbf{T}(i,j)$. For instance,

$$\mathbf{T}(1,1) = \frac{det[\mathbf{C}_{jk+1} \quad \mathbf{C}_{j2} \quad \mathbf{C}_{j3} \quad \cdots \quad \mathbf{C}_{jn}]}{det[\mathbf{C}_{j1} \quad \mathbf{C}_{j2} \quad \cdots \quad \mathbf{C}_{jk}]} \tag{8}$$

By definition of $\mathbf{J}_s$, the determinant in the denominator is at least as large as the determinant in the numerator in magnitude. The same argument applies for all i, j

$$\mathbf{T}(2,1) = \frac{det[\mathbf{C}_{j1} \quad \mathbf{C}_{jk+1} \quad \mathbf{C}_{j2} \quad \cdots \quad \mathbf{C}_{jn}]}{det[\mathbf{C}_{j1} \quad \mathbf{C}_{j2} \quad \cdots \quad \mathbf{C}_{jk}]} \tag{9}$$

Case 2: Let m $\geq$ k. The $\mathbf{A}$ admits some factorization

$$\underset{m \times n}{\mathbf{A}} = \underset{m \times k}{\mathbf{E}} \; \underset{k \times n}{\mathbf{F}}. \tag{10}$$

We show that $\mathbf{F}$ admits a factorization

$$\mathbf{F} = \mathbf{F}(:, \mathbf{J}_s)\,\mathbf{Z}. \tag{11}$$

For $\mathbf{Z}$ that satisfies the criteria in Theorem 1. Combining Equations 10 and 11

$$\mathbf{A} = \mathbf{E}\mathbf{F}(:, \mathbf{J}_s)\,\mathbf{Z} \tag{12}$$

Restricting Equation 12 to $\mathbf{J}_s$,

$$\mathbf{A}(:, \mathbf{J}_s) = \mathbf{E}\mathbf{F}(:, \mathbf{J}_s)\,\mathbf{Z}(:, \mathbf{J}_s) \tag{13}$$

Note that $\mathbf{Z}(:, \mathbf{J}_s) = \mathbf{I}_k$ by construction. Combining Equations 12 and 13,

$$\mathbf{A} = \mathbf{A}(:, \mathbf{J}_s)\,\mathbf{Z} \tag{14}$$

2. Deterministic Methods for Computing the ID

2.1. **Option 1.** Do as in the proof. This is extremely expensive, but it leads to an optimal result

2.2. **Option 2.** Use the CPQR from the last lecture.

$$\mathbf{A} \approx \mathbf{Q}_1\left[\mathbf{R}_{1,1} \quad \mathbf{R}_{1,2}\right]\mathbf{P}^* = \mathbf{Q}\mathbf{R}_{1,1}\begin{bmatrix}\mathbf{I} & \mathbf{R}_{1,1}^{-1}\mathbf{R}_{1,2}\end{bmatrix}\mathbf{P}^*$$

With $\mathbf{C} = \mathbf{Q}\mathbf{R}_{1,1}$ and $\mathbf{Z} = \begin{bmatrix}\mathbf{I} & \mathbf{R}_{1,1}^{-1}\mathbf{R}_{1,2}\end{bmatrix}\mathbf{P}^*$. This method is very computationally efficient. You can use the standard software to accomplish this method. In practice, elements of $\mathbf{Z}$ are more or less bounded in magnitude by 1. Artificial counterexamples exist.

2.3. **Option 3. Rank Revealing QR** (recommend)
Let $\nu > 0$ be a positive number. There exist pivoting strategies that find, in polynomial time, an index set $\mathbf{J}_s$ such that $\|\mathbf{T}(i,j)\| \leq 1+\nu$. They also attain a "goodish" rank-k approximation $\|\mathbf{A} - \mathbf{Q}_1[\mathbf{R}_{11}\mathbf{R}_{12}]\,\mathbf{P}^*\| = \|\mathbf{R}_{22}\| \approx \sigma_{k+1} \leq p(k,n)\,\sigma_{k+1}$.

3. Randomized ID

First Recall the RSVD

$$\begin{aligned}
\mathbf{G} &= rand(n, k+p)\\
\mathbf{Y} &= \mathbf{A}\mathbf{G}\\
\mathbf{Q} &= orth(\mathbf{Y})\\
\mathbf{B} &= \mathbf{Q}^*\mathbf{A}\\
\mathbf{B} &= \hat{\mathbf{U}}\mathbf{D}\mathbf{V}^*\\
\mathbf{U} &= \mathbf{Q}\hat{\mathbf{U}}\\
\mathbf{A} &\approx \mathbf{U}\mathbf{D}\mathbf{V}^*
\end{aligned}$$

Assume that $\mathbf{A}$ has exact rank k.

$$\begin{aligned}
\mathbf{G} &= rand(n, k)\\
\mathbf{Y} &= \mathbf{A}\mathbf{G}\\
[\mathbf{I}_s, \mathbf{X}] &= IDrow(\mathbf{Y}, k)
\end{aligned}$$

$$\mathbf{Y} = \mathbf{X}\mathbf{Y}(\mathbf{I}_s, :) \tag{15}$$

Automatically $\mathbf{A} = \mathbf{X}\mathbf{A}(\mathbf{I}_s, :)$. With probability 1, we know that for some $\mathbf{F}$

$$\mathbf{A} = \mathbf{Y}\mathbf{F} \tag{16}$$

Combining Equation 16 and 15,

$$\mathbf{A} = \mathbf{X}\mathbf{Y}\left(\mathbf{I}_s, :\right)\mathbf{F} \tag{17}$$

Restricting Equation 17,

$$\mathbf{A}\left(\mathbf{I}_s, :\right) = \mathbf{X}\left(\mathbf{I}_s, :\right)\mathbf{Y}\left(\mathbf{I}_s, :\right)\mathbf{F} \tag{18}$$

Combining Equation 18 and 17,

$$\mathbf{A} = \mathbf{X}\mathbf{A}\left(\mathbf{I}_s, :\right) \tag{19}$$