Fast Matrix Algorithms for Data Analytics: Problem Set 2

1. Let $\mathbf{A} \in \mathbb{R}^{m\times n}$, J be an $n \times 1$ permutation vector, and $J_s = J(1:k)$ for some $k \in \mathbb{N}$. Now suppose
$$\underset{m\times n}{\mathbf{A}} = \underset{m\times k}{\mathbf{E}} \; \underset{k\times n}{\mathbf{F}}, \tag{1}$$
and suppose that for some matrix $\mathbf{T}$ of size $k \times (n-k)$, it holds that
$$\mathbf{F}(:,J) = \mathbf{F}(:,J_s)\begin{bmatrix}\mathbf{I}_k & \mathbf{T}\end{bmatrix},$$
where $\mathbf{I}_k$ is the $k\times k$ identity matrix. Prove that
$$\mathbf{A}(:,J) = \mathbf{A}(:,J_s)\begin{bmatrix}\mathbf{I}_k & \mathbf{T}\end{bmatrix}.$$

2. Suppose that $\mathbf{A}$ is an $m \times n$ matrix of precise rank k. Moreover, suppose that you have available a double-sided ID of $\mathbf{A}$ of the form
$$\underset{m\times n}{\mathbf{A}} = \underset{m\times k}{\mathbf{X}} \; \underset{k\times k}{\mathbf{A}_{\mathrm{s}}} \; \underset{k\times n}{\mathbf{Z}}, \tag{2}$$
where $\mathbf{A}_{\mathrm{s}} = \mathbf{A}(I_{\mathrm{s}}, J_{\mathrm{s}})$ for some index vectors I_{s} and J_{s}. Recall from the definition of the double-sided ID that $\mathbf{X}$ and $\mathbf{Z}$ must contain the $k \times k$ identity matrix $\mathbf{I}_k$ as a submatrix, so that $\mathbf{X}(I_{\mathrm{s}},:) = \mathbf{I}_k$ and $\mathbf{Z}(:,J_{\mathrm{s}}) = \mathbf{I}_k$.

   (a) Prove that $\mathbf{A}_{\mathrm{s}}$ must be of full rank.

   (b) Set $\mathbf{U} = \mathbf{A}_{\mathrm{s}}^{-1}$, so that (2) can be written
$$\underset{m\times n}{\mathbf{A}} = \underset{m\times k}{\mathbf{X}\mathbf{A}_{\mathrm{s}}} \; \underset{k\times k}{\mathbf{U}} \; \underset{k\times n}{\mathbf{A}_{\mathrm{s}}\mathbf{Z}}. \tag{3}$$
   Prove that $\mathbf{X}\mathbf{A}_{\mathrm{s}} = \mathbf{A}(:,J_{\mathrm{s}})$ and $\mathbf{A}_{\mathrm{s}}\mathbf{Z} = \mathbf{A}(I_{\mathrm{s}},:)$ so that (3) is a CUR factorization.

3. Suppose that $\mathbf{A}$ 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

$$\mathbf{C} = \mathbf{A}(:, J_s), \quad \mathbf{R} = \mathbf{A}(I_s, :)$$

hold k columns/rows that span the column/row space of $\mathbf{A}$. Then

$$\mathbf{A} \approx \mathbf{C}\mathbf{C}^{\dagger}\mathbf{A}\mathbf{R}^{\dagger}\mathbf{R},$$

and the optimal choice for the "U" factor in the CUR decomposition is

$$\mathbf{U} = \mathbf{C}^{\dagger}\mathbf{A}\mathbf{R}^{\dagger}.$$

Set $\mathbf{X} = \mathbf{C}\mathbf{C}^{\dagger}$.

(a) Suppose that $\mathbf{C}$ has the SVD $\mathbf{C} = \mathbf{W}\mathbf{D}\mathbf{V}^*$. Prove that $\mathbf{X} = \mathbf{W}\mathbf{W}^*$.

(b) Suppose that $\mathbf{C}$ has the QR factorization $\mathbf{C}\mathbf{P} = \mathbf{Q}\mathbf{S}$. Prove that $\mathbf{X} = \mathbf{Q}\mathbf{Q}^*$.

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

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

4. Let $\mathbf{A} \in \mathbb{R}^{m\times n}$ have rank exactly k. In this problem, we will prove that $\mathbf{A}$ admits a factorization $\mathbf{A} = \mathbf{A}(:, J_s)\mathbf{Z}$, where $\mathbf{A}(:, J_s) \in \mathbb{R}^{m\times k}$ and $\mathbf{Z} \in \mathbb{R}^{k\times n}$ such that $\mathbf{Z}(:, J_s) = \mathbf{I}_k$ and $\max_{i,j} |\mathbf{Z}(i, j)| \leq 1$.

   (a) **case 1:** $m = k$.

      i. Pick a permutation vector J_s such that $|\det(\mathbf{A}(:, J_s))|$ is maximized, and let J_r denote the remaining indices so that $[J_s \quad J_r]$ is some permutation of the vector $[1\ 2 \quad \cdots \quad n]$. Then we have that

$$\mathbf{A}(:, [J_s \quad J_r]) = [\mathbf{A}(:, J_s) \quad \mathbf{A}(:, J_r)]$$

      can be written as $\mathbf{AP}$ for some permutation matrix $\mathbf{P}$. Find an interpolative decomposition $\mathbf{A} = \mathbf{CZ}$ of $\mathbf{A}$, where the columns of $\mathbf{C}$ are some of the columns of $\mathbf{A}$. $\mathbf{C}$ and $\mathbf{Z}$ should be in terms of $\mathbf{A}(:, J_s), \mathbf{A}(:, J_r), \mathbf{P}$, and the identity matrix $\mathbf{I}$.

      ii. Consider the matrix $\mathbf{T} = \mathbf{A}(:, J_s)^{-1}\mathbf{A}(:, J_r)$. If we can show that

$$\max_{i,j} |\mathbf{T}(i, j)| \leq 1, \tag{4}$$

      then we will be done with the case $m = k$ (why?). Find a way to show (4) by applying Cramer's Rule to our definition of $\mathbf{T}$.

      Cramer's Rule: Consider the linear system $\mathbf{Ax} = \mathbf{b}$. The i-th entry of the solution $\mathbf{x}$ is given by

$$x_i = \frac{\det(\mathbf{A}_i)}{\det(\mathbf{A})},$$

      where $\mathbf{A}_i$ is matrix formed by replacing the i-th column of $\mathbf{A}$ with $\mathbf{b}$.

   (b) **case 2:** $m \geq k$.
   Then $\mathbf{A}$ admits a factorization $\mathbf{A} = \mathbf{EF}$, where $\mathbf{E}$ is $m \times k$ and $\mathbf{F}$ is $k \times n$. Apply case 1 to $\mathbf{F}$ to show the result for this case (something we proved in a previous problem may help, too...).