Fast Matrix Algorithms for Data Analytics: Problem Set 1

1. Assume that for $\mathbf{A} \in \mathbb{R}^{m\times n}$ and $\mathbf{Q} \in \mathbb{R}^{m\times k}$ where $\mathbf{Q}$ has orthonormal columns, $\text{range}(\mathbf{Q}) = \text{range}(\mathbf{A})$. Prove that $\mathbf{A} = \mathbf{Q}\mathbf{Q}^*\mathbf{A}$.

   Hint: A linear operator $P : X \to Y$ between two vector spaces X and Y is a projection iff it satisfies $P^2 = P$. Projections satisfy the property that for all $x \in \text{range}(P)$, $Px = x$. (this follows from the definition; if you don't see why, prove it!) If you don't see how to proceed with the proof, try an approach that takes advantage of this information.

2. (a) Let $\mathbf{A}$ be an $m \times n$ matrix, set $p = \min(m, n)$, and suppose that the singular value decomposition of $\mathbf{A}$ takes the form

$$\underset{m\times n}{\mathbf{A}} = \underset{m\times p}{\mathbf{U}} \; \underset{p\times p}{\mathbf{D}} \; \underset{p\times n.}{\mathbf{V}^*} \tag{1}$$

   Recall the definition of the spectral norm of $\mathbf{A}$:

$$\|\mathbf{A}\| = \sup_{\mathbf{x}\neq 0} \frac{\|\mathbf{A}\mathbf{x}\|_2}{\|\mathbf{x}\|_2} = \sup_{\|\mathbf{x}\|_2=1} \|\mathbf{A}\mathbf{x}\|_2.$$

   Let k be an integer such that $1 \leq k < p$ and let $\mathbf{A}_k$ denote the truncation of the SVD to the first k terms:

$$\mathbf{A}_k = \mathbf{U}(:, 1 : k)\mathbf{D}(1 : k, 1 : k)\mathbf{V}(:, 1 : k)^*.$$

   Prove directly from the definition of the spectral norm that

$$\|\mathbf{A} - \mathbf{A}_k\| = \sigma_{k+1}. \tag{2}$$

   (b) In phase A of the `RSVD` algorithm, we seek a matrix $\mathbf{Q} \in \mathbb{R}^{m\times k}$ with orthonormal columns such that $\mathbf{A} \approx \mathbf{Q}\mathbf{Q}^*\mathbf{A}$. Since $\text{rank}(\mathbf{Q}\mathbf{Q}^*\mathbf{A}) \leq k$, the Eckart-Young Theorem assures us that

$$\inf_{\mathbf{Q}\in\mathbb{R}^{m\times k}} \|\mathbf{A} - \mathbf{Q}\mathbf{Q}^*\mathbf{A}\| \geq \sigma_{k+1}.$$

   Show that we can achieve this bound by choosing $\mathbf{Q} = \begin{bmatrix} \mathbf{u}_1 & \mathbf{u}_2 & \cdots & \mathbf{u}_k \end{bmatrix}$, where $\{\mathbf{u}_i\}_{i=1}^k$ are the k leading left singular vectors of $\mathbf{A}$. That is, show that for such a $\mathbf{Q}$, we have

$$\|\mathbf{A} - \mathbf{Q}\mathbf{Q}^*\mathbf{A}\| = \sigma_{k+1}.$$

3. Suppose $\mathbf{A}$ is a real symmetric $n \times n$ matrix with eigenpairs $\{\lambda_j, \mathbf{v}_j\}_{j=1}^n$, ordered so that $|\lambda_1| \geq |\lambda_2| \geq \ldots \geq |\lambda_n|$. Define a sequence of vectors $\mathbf{x}_p = \mathbf{A}^p \mathbf{v}_0$, where $\mathbf{v}_0$ is an $n \times 1$ random vector whose entries are drawn independently from a standard Gaussian distribution.

   (a) Set $\beta = |\lambda_2|/|\lambda_1|$ and $\mathbf{y}_p = (1/\|\mathbf{x}_p\|)\mathbf{x}_p$. Assume $\lambda_1 = 1$ and $\beta < 1$. Prove that as $p \to \infty$, the vectors $\{\mathbf{y}_p\}$ converge either to $\mathbf{v}_1$ or $-\mathbf{v}_1$.

   (b) What is the speed of convergence of $\{\mathbf{y}_p\}$?

   (c) Assume again that $\beta < 1$, but now drop the assumption that $\lambda_1 = 1$. Prove that your answers in (a) and (b) are still correct, with the exception that if $\lambda_1 < 0$, then it is the vector $(-1)^p \mathbf{y}_p$ that converges instead.

4. Consider the "single pass algorithm" for a non-Hermitian $m \times n$ matrix $\mathbf{A}$, the essentials of which are reiterated below. Suppose our matrix $\mathbf{T}$ is such that $\mathbf{Q}^*\mathbf{Y} = \mathbf{T}(\mathbf{W}^*\mathbf{G})$ and $\mathbf{W}^*\mathbf{Z} = \mathbf{T}^*(\mathbf{Q}^*\mathbf{H})$ hold *exactly* (this is not usually the case in practice!). Show that in this case, the output of the algorithm is exact, *i.e.* $\mathbf{A} = \mathbf{U}\mathbf{D}\mathbf{V}^*$.

"Single Pass" RSVD

- Generate random matrices $\mathbf{G} \in \mathbb{R}^{n \times \ell}$ and $\mathbf{H} \in \mathbb{R}^{m \times \ell}$. For this problem, choose $\ell = \min(m, n)$, though in practice we choose $\ell < \min(m, n)$.
- Compute sample matrices $\mathbf{Y} = \mathbf{A}\mathbf{G}$ and $\mathbf{Z} = \mathbf{A}^*\mathbf{H}$.
- Find ON matrices $\mathbf{Q}$ and $\mathbf{W}$ such that $\mathbf{Y} = \mathbf{Q}\mathbf{Q}^*\mathbf{Y}$ and $\mathbf{Z} = \mathbf{W}\mathbf{W}^*\mathbf{Z}$.
- Solve for $\mathbf{T}$ the linear systems $\mathbf{Q}^*\mathbf{Y} = \mathbf{T}(\mathbf{W}^*\mathbf{G})$ and $\mathbf{W}^*\mathbf{Z} = \mathbf{T}^*(\mathbf{Q}^*\mathbf{H})$.
- Factor $\mathbf{T}$ so that $\mathbf{T} = \hat{\mathbf{U}}\mathbf{D}\hat{\mathbf{V}}^*$.
- Form $\mathbf{U} = \mathbf{Q}\hat{\mathbf{U}}$ and $\mathbf{V} = \mathbf{W}\hat{\mathbf{V}}$.