1. Finding $\mathbf{Q}_j$ such that $\|\mathbf{Q}_j\mathbf{Q}_j^*\mathbf{A} - \mathbf{A}\| < \epsilon$ is guaranteed

A shortcoming of the last algorithm discussed in the previous lecture is that the resulting matrix $\mathbf{Q}$ is not *guaranteed* to satisfy $\|\mathbf{Q}\mathbf{Q}^*\mathbf{A} - \mathbf{A}\| < \epsilon$; it merely achieves this requirement with a high degree of probability. In the case that we require certainty in our result, another approach is required. One such approach is given by the following algorithm:

- **Algorithm 1**
  - Given inputs $\mathbf{A}$ and $\epsilon > 0$, set $\mathbf{A}_0 = \mathbf{A}, \mathbf{B}_0 = [\quad], \mathbf{Q}_0 = [\quad]$.
  - for $j = 1, 2, \ldots$
    * Choose an i.i.d. $\mathbf{w} \in \mathbb{R}^n$, set $\mathbf{q}_j = \mathbf{A}\mathbf{w}/\|\mathbf{A}\mathbf{w}\|$.
    * Set $\mathbf{b}_j = \mathbf{q}_j^*\mathbf{A}_{j-1}$.
    * Set $\mathbf{Q}_j = [\mathbf{Q}_{j-1}\ \ \mathbf{q}_j]$.
    * Set $\mathbf{B}_j = \begin{bmatrix} \mathbf{B}_{j-1} \\ \mathbf{b}_j \end{bmatrix}$.
    * Set $\mathbf{A}_j = \mathbf{A}_{j-1} - \mathbf{q}_j\mathbf{b}_j = \mathbf{A}_{j-1} - \mathbf{q}_j\mathbf{q}_j^*\mathbf{A}_{j-1}$.
    * If $\|\mathbf{A}_j\| < \epsilon$, exit.

We will next show that the algorithm "works," that is, we will show that $\mathbf{A}_j = \mathbf{Q}_j\mathbf{Q}_j^*\mathbf{A} - \mathbf{A}$, so when we terminate the routine (i.e. when $\|\mathbf{A}_j\| < \epsilon$), we also have $\|\mathbf{Q}_j\mathbf{Q}_j^*\mathbf{A} - \mathbf{A}\| < \epsilon$. We will do this by induction.

Iteration 1. Here we have $\mathbf{Q}_1 = [\mathbf{q}_1]$ and $\mathbf{A}_0 = \mathbf{A}$, so

$$\mathbf{A}_1 = \mathbf{A}_0 - \mathbf{q}_1\mathbf{q}_1^*\mathbf{A}_0 = (\mathbf{I} - \mathbf{Q}_1\mathbf{Q}^*)\mathbf{A}.$$

Iteration j/Induction Step. Assume we have $\mathbf{A}_j = (\mathbf{I} - \mathbf{Q}_j\mathbf{Q}_j^*)\mathbf{A}$.

$$\mathbf{A}_{j+1} = \mathbf{A}_j - \mathbf{q}_{j+1}\mathbf{q}^*\mathbf{A}_j = (\mathbf{A} - \mathbf{Q}_j\mathbf{Q}_j^*\mathbf{A}) - \mathbf{q}_{j+1}\mathbf{q}_{j+1}^*(\mathbf{A} - \mathbf{Q}_j\mathbf{Q}_j^*\mathbf{A}).$$

Now since $\mathbf{q}_{j+1}$ is orthogonal to the columns of $\mathbf{Q}_j$, $\mathbf{q}_{j+1}^*\mathbf{Q}_j = \mathbf{0}$. Therefore

$$\begin{aligned}(\mathbf{A} - \mathbf{Q}_j\mathbf{Q}_j^*\mathbf{A}) - \mathbf{q}_{j+1}\mathbf{q}_{j+1}^*(\mathbf{A} - \mathbf{Q}_j\mathbf{Q}_j^*\mathbf{A}) &= \mathbf{A} - \mathbf{Q}_j\mathbf{Q}_j^*\mathbf{A} - \mathbf{q}_{j+1}\mathbf{q}_{j+1}^*\mathbf{A} \\ &= (\mathbf{I} - (\mathbf{Q}_j\mathbf{Q}_j^* + \mathbf{q}_{j+1}\mathbf{q}_{j+1}^*))\mathbf{A}.\end{aligned}$$

Then since

$$\mathbf{Q}_{j+1}\mathbf{Q}_{j+1}^* = \begin{bmatrix} \mathbf{Q}_j \\ \mathbf{q}_{j+1} \end{bmatrix} \begin{bmatrix} \mathbf{Q}_j^* & \mathbf{q}_{j+1}^* \end{bmatrix} = \mathbf{Q}_j\mathbf{Q}_j^* + \mathbf{q}_{j+1}\mathbf{q}_{j+1}^*,$$

we have that

$$\mathbf{A}_{j+1} = (\mathbf{I} - \mathbf{Q}_{j+1}\mathbf{Q}_{j+1}^*)\mathbf{A}. \tag{1}$$

2. BLOCKING ALGORITHM 1

One disadvantage of Algorithm 1 is that the matrix $\mathbf{Q}$ is updated one vector at a time, so the algorithm relies mostly on matrix-vector products rather than the more efficient matrix-matrix products. The following modification of Algorithm 1 remedies this problem.

- **Algorithm 2**
  - Set $\mathbf{A}_0 = \mathbf{A}, \mathbf{B}_0 = [\ ], \mathbf{Q}_0 = [\ ]$.
  - for $j = 1, 2, \ldots$
    - Set $\hat{\mathbf{Q}}_j = \mathtt{qr}(\mathbf{A}_{j-1}\mathbf{\Omega}_{j-1}, 0)$, where $\Omega \in \mathbb{R}^{n\times b}$ matrix whose entries are i.i.d. Gaussian random variables, where b is some block size.
    - Set $\hat{\mathbf{B}}_j = \hat{\mathbf{Q}}_j^*\mathbf{A}_{j-1}$.
    - Set $\mathbf{A}_j = \mathbf{A}_{j-1} - \hat{\mathbf{Q}}_j\hat{\mathbf{B}}_j$.
    - Set $\mathbf{B}_j = \begin{bmatrix}\mathbf{B}_{j-1}^* & \hat{\mathbf{B}}_j^*\end{bmatrix}, \mathbf{Q}_j = \begin{bmatrix}\mathbf{Q}_{j-1} & \hat{\mathbf{Q}}_j\end{bmatrix}$.
    - If $\|\mathbf{A}_j\| < \epsilon$, exit.

We would again like to show that the algorithm gives the expected result, which means we must show that $\mathbf{A}_j = (\mathbf{I} - \mathbf{Q}_j\mathbf{Q}_j^*)\mathbf{A}$.

We will first demonstrate that $\hat{\mathbf{Q}}_j^*\hat{\mathbf{Q}}_k = \mathbf{0}$ for $k = 1, \ldots, j-1$. The desired result will then follow. We have that $\mathrm{Ran}(\hat{\mathbf{Q}}_j) = \mathrm{Ran}(\mathbf{A}_{j-1}\mathbf{\Omega}) \subseteq \mathrm{Ran}(\mathbf{A}_{j-1}) = \mathrm{Ran}((\mathbf{I} - \sum_{k=1}^{j-1}\hat{\mathbf{Q}}_k\hat{\mathbf{Q}}_k^*)\mathbf{A}) \subseteq \mathrm{Ran}(\mathbf{I} - \sum_{k=1}^{j-1}\hat{\mathbf{Q}}_k\hat{\mathbf{Q}}_k^*) = \mathrm{Ran}(\sum_{k=1}^{j-1}\hat{\mathbf{Q}}_k\hat{\mathbf{Q}}_k^*)^\perp$. Thus $\hat{\mathbf{Q}}_j^*\hat{\mathbf{Q}}_k = \mathbf{0}$ for $k = 1, \ldots, j-1$.

Now we have that

$$\begin{aligned}\mathbf{A}_j &= (\mathbf{I} - \hat{\mathbf{Q}}_j\hat{\mathbf{Q}}_j^*)\mathbf{A}_{j-1} \\ &= (\mathbf{I} - \hat{\mathbf{Q}}_j\hat{\mathbf{Q}}_j^*)(\mathbf{I} - \hat{\mathbf{Q}}_{j-1}\hat{\mathbf{Q}}_{j-1}^*)\mathbf{A}_{j-2} \\ &= \ldots = (\mathbf{I} - \hat{\mathbf{Q}}_j\hat{\mathbf{Q}}_j^* - \hat{\mathbf{Q}}_{j-1}\hat{\mathbf{Q}}_{j-1}^* - \ldots - \hat{\mathbf{Q}}_1\hat{\mathbf{Q}}_1^*)\mathbf{A} \\ &= (\mathbf{I} - \mathbf{Q}_j\mathbf{Q}_j^*)\mathbf{A}.\end{aligned}$$