Justin Hexem
AMATH 584

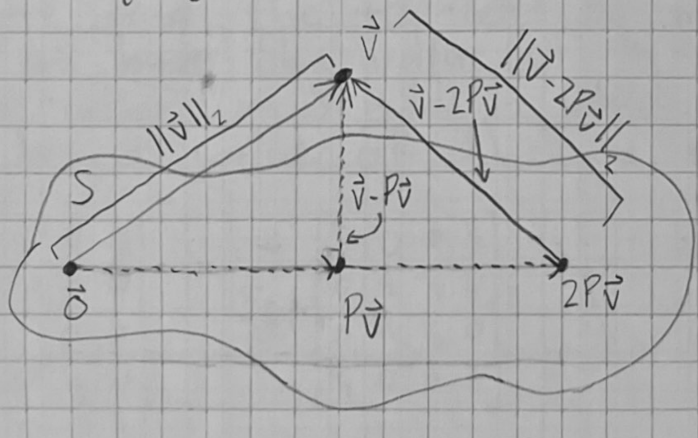
Homework #3

6.1.) Proof: Let P be an orthogonal projector ($P^2 = P$ & $P = P^*$). Show $(I-2P)$ is unitary (Show $(I-2P)^* = (I-2P)^{-1}$). We see that

$$(I-2P)^*(I-2P) = (I^* - (2P)^*)(I-2P) = (I-2P^*)(I-2P)$$
$$= (I-2P)(I-2P) = I^2 - 4P + 4P^2$$
$$= I - 4P + 4P = I$$

Since $(I-2P)^*(I-2P) = I$ we know $(I-2P)^* = (I-2P)^{-1}$. Thus $(I-2P)$ must be unitary. $\boxed{\times}$

We can draw a geometric interpretation that helps to illustrate this in the following manner. Let S be a subspace of $\mathbb{C}^m$, let $v \in \mathbb{C}^m$, and let P be the orthogonal projector onto S.



This picture shows the projection of $\vec{v}$ onto S, the stretching of $P\vec{v}$ to $2P\vec{v}$, and the difference $\vec{v} - 2P\vec{v}$. The proof states that $I-2P$ is unitary, so in the picture we see that $\|(I-2P)\vec{v}\|_2 = \|\vec{v} - 2P\vec{v}\|_2 = \|\vec{v}\|_2$ since norms are preserved by unitary matrices under the 2-norm.

Scanned with CamScanner

6.2.) Let $E = \frac{1}{2}Ix + \frac{1}{2}Fx$ where $F((x_1, x_2, \ldots, x_m)^*) = (x_m, x_{m-1}, \ldots, x_1)^*$.
First note that F is the matrix with ones down the bottom-left to top-right diagonal and zeros everywhere else. This means we can trivially see that $F^* = F$. Let $x \in \mathbb{C}^m$.
Then

$$Ex = \frac{1}{2}x + \frac{1}{2}Fx$$

$$E^2x = E(Ex) = E\left(\frac{1}{2}x + \frac{1}{2}Fx\right) = E\left(\frac{1}{2}x\right) + E\left(\frac{1}{2}Fx\right)$$

$$= \frac{1}{2}\left(\frac{1}{2}x + \frac{1}{2}Fx\right) + \frac{1}{2}\left(\frac{1}{2}Fx + \frac{1}{2}F^2x\right) = \frac{1}{4}x + \frac{1}{4}Fx + \frac{1}{4}Fx + \frac{1}{4}F^2x$$

$$= \frac{1}{4}x + \frac{1}{2}Fx + \frac{1}{4}x = \frac{1}{2}x + \frac{1}{2}Fx = Ex$$

$$E^*x = \left(\frac{1}{2}I + \frac{1}{2}F\right)^*x = \left(\frac{1}{2}I\right)^*x + \left(\frac{1}{2}F\right)^*x = \frac{1}{2}I^*x + \frac{1}{2}F^*x$$

$$= \frac{1}{2}Ix + \frac{1}{2}Fx = \frac{1}{2}x + \frac{1}{2}Fx = Ex$$

Since $E^2 = E$ and $E^* = E$ we conclude that E is an orthonormal projector. We see that E has entries of all zero except for on both the bottom-left to top-right and the top-left to bottom-right diagonals which all have the value of $\frac{1}{2}$ (except possibly in the case where m is odd, in which case the center entry has a value of 1). Two examples are given:

$$E_5 = \begin{bmatrix} \frac{1}{2} & 0 & 0 & 0 & \frac{1}{2} \\ 0 & \frac{1}{2} & 0 & \frac{1}{2} & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & \frac{1}{2} & 0 & \frac{1}{2} & 0 \\ \frac{1}{2} & 0 & 0 & 0 & \frac{1}{2} \end{bmatrix} \qquad E_4 = \begin{bmatrix} \frac{1}{2} & 0 & 0 & \frac{1}{2} \\ 0 & \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & \frac{1}{2} & \frac{1}{2} & 0 \\ \frac{1}{2} & 0 & 0 & \frac{1}{2} \end{bmatrix}$$

7.8.) a.) Proof: Let $A \in \mathbb{C}^{m \times n}$ and $A = \hat{Q}\hat{R}$ be a reduced QR factorization. Show that A has rank n if and only if all the diagonal entries of $\hat{R}$ are nonzero.

($\Rightarrow$) Assume A has rank n. Then this implies both $\hat{Q}$ & $\hat{R}$ also have rank of at least n. Since both $\hat{Q}$ & $\hat{R}$ have n columns, this means $\text{rank}(\hat{Q}) = \text{rank}(\hat{R}) = n$. Since $\hat{R}$ has rank n and $\hat{R} \in \mathbb{C}^{n \times n}$, we know $\det(\hat{R}) \neq 0$. Since $\det(\hat{R}) = \prod_{k=1}^{n} r_{kk} = 0$, we conclude that $r_{kk} \neq 0$ for all $k \in \{1, 2, \dots, n\}$. $\square$

($\Leftarrow$) Assume $r_{kk} \neq 0$ for all $k \in \{1, 2, \dots, n\}$. Then we know $\det(\hat{R}) = \prod_{k=1}^{n} r_{kk} \neq 0$, so $\hat{R}$ has rank n. We also know that every column of $\hat{Q}$ is orthogonal to every other column of $\hat{Q}$, so all n columns of $\hat{Q}$ are linearly independent. This means $\hat{Q}$ has rank n as well. Since A is the product of two rank n matrices, A must also have rank n. $\square$

Therefore A has rank n if and only if all the diagonal entries of $\hat{R}$ are nonzero. $\boxtimes$

b.) Proof: Let $A \in \mathbb{C}^{m \times n}$ and let $A = \hat{Q}\hat{R}$ be a reduced QR factorization. Suppose $\hat{R}$ has k nonzero diagonal entries for some k with $0 \leq k < n$. Show $\text{rank}(A) \geq k$. We will proceed by observing what happens when we set some nonzero diagonal entry r_{jj} to zero for some $j \in \{1, 2, \dots, n\}$. We have four cases.

Case 1: $r_{(j-1)(j-1)} \neq 0$ & $r_{(j+1)(j+1)} \neq 0$.
If this is the case, then setting $r_{jj} = 0$ would mean $a_j \in \langle q_1, q_2, \dots, q_{j-1} \rangle = \langle a_1, a_2, \dots, a_{j-1} \rangle$, which implies $\text{rank}(A)$ decreases by 1 as k also decreases by 1.

Case 2: $r_{(j-1)(j-1)} = r_{(j+1)(j+1)} = 0$.
If this is the case then $a_{j-1} \in \langle q_1, q_2, \dots, q_{j-2} \rangle = \langle a_1, a_2, \dots, a_{j-2} \rangle$ and $a_{j+1} \in \langle q_1, q_2, \dots, a_j \rangle = \langle a_1, a_2, \dots, a_j \rangle$. When we set $r_{jj} = 0$ we see that a_j moves from $\langle a_1, a_2, \dots, a_j \rangle$ to $\langle a_1, a_2, \dots, a_{j-1} \rangle$ so $\text{rank}(A)$ increases by 1 as k decreases by 1.

Case 3: $r_{(j-1)(j-1)} = 0$ & $r_{(j+1)(j+1)} \neq 0$
If this is the case, then $a_{j-1} \in \langle a_1, a_2, \dots, a_{j-2} \rangle$ and $a_{j+1} \in \langle a_1, a_2, \dots, a_{j+1} \rangle$. This means setting $r_{jj} = 0$ moves a_j from $\langle a_1, a_2, \dots, a_j \rangle$ to $\langle a_1, a_2, \dots, a_{j-1} \rangle$. Since both of these column spaces are distinct from those that a_{j-1} and a_{j+1} reside in, rank(A) remains unchanged as k decreases by 1.

Case 4: $r_{(j-1)(j-1)} \neq 0$ & $r_{(j+1)(j+1)} = 0$.
If this is the case then $a_{j-1} \in \langle a_1, a_2, \dots, a_{j-1} \rangle$ and $a_{j+1} \in \langle a_1, a_2, \dots, a_j \rangle$. This means setting $r_{jj} = 0$ moves a_j from $\langle a_1, a_2, \dots, a_j \rangle$ to $\langle a_1, a_2, \dots, a_{j-1} \rangle$. This means a_j did not change the number of distinct nested column spaces of A, so rank(A) remains unchanged as k decreases by 1.

Therefore as k decreases by 1, rank(A) cannot decrease by more than 1 in all cases so rank(A) must be at least as large as k. Cases 2-4 show that rank(A) can also be larger than k, so we conclude that rank(A) ≥ k. ☒

3.) Proof: Let $A \in \mathbb{C}^{m \times n}$, $m \geq n$, and suppose $\text{rank}(A) = k < n$. Show there exists a factorization $A = Q_k R_k P^T$, where Q_k is an $m \times k$ matrix with orthonormal columns, R_k is an upper triangular $k \times n$ matrix, and P is a permutation matrix. We will prove this by construction. First choose a permutation matrix P such that the first k columns of AP are linearly independent. This means if a_j is the jth column of AP where $j < k$, then $\langle a_1, a_2, \dots, a_k \rangle = \text{range}(AP) = \text{range}(A)$. By 7.1 we can take the reduced QR factorization of AP as $AP = \hat{Q}\hat{R}$. We know $\langle q_1, q_2, \dots, q_j \rangle = \langle a_1, a_2, \dots, a_j \rangle$ for all $j \leq k$, so $\langle q_1, q_2, \dots, q_k \rangle = \text{range}(A)$. This means $a_{k+1}, a_{k+2}, \dots, a_n \in \langle q_1, q_2, \dots, q_k \rangle$ so we see that for $j \in \{k+1, k+2, \dots, n\}$ we know $a_j = r_{1j} q_1 + \dots + r_{kj} q_k$. This means $r_{xy} = 0$ for all $k < x \leq m$ and $1 \leq y \leq n$, so the product $\hat{Q}\hat{R}$ does not use any of the last $n-k$ columns at all since the zero entries cancel them out. Let us form the matrix Q_k which consists of the first k columns of $\hat{Q}$. We can also form the matrix R_k which consists of the first k rows of $\hat{R}$ that omits the last $m-k$ rows of zeros. Since we have not lost any information in constructing Q_k and R_k we see that $\hat{Q}\hat{R} = Q_k R_k$. Thus $AP = Q_k R_k$ which implies that $A = Q_k R_k P^T$. Since the matrices Q_k, R_k, and P^T satisfy the conditions stated above we conclude that such a factorization of A exists. ☒

4.) a.) Let $A \in \mathbb{C}^{m \times m}$ be a rank k matrix, and suppose we know and can access factors $X, Y \in \mathbb{C}^{m \times k}$ such that $A = XY^*$. We will create an algorithm that computes the skinny svd of A from X & Y in $\mathcal{O}(mk^2 + k^3)$ floating point operations

[U, S, V] = LRsvd(X, Y)

Skinny QR decompose X & Y ($\mathcal{O}(mk^2)$ flops)

$X = Qkx\ Rkx\ Px \quad , \quad Y = Qky\ Rky\ Py$

$$
\begin{aligned}
\% \quad A = XY^* &= (Qkx \cdot Rkx \cdot Px^T)(Qky \cdot Rky \cdot Py^T)^* \\
\% \quad &= Qkx \cdot Rkx \cdot Px^T \cdot Py \cdot Rky^* \cdot Qky^* \\
\% \quad &= Qkx \cdot (Rkx \cdot Px^T \cdot Py \cdot Rky^*) \cdot Qky^* \\
\% \quad &= Qkx \cdot ((Rkx \cdot Px^T) \cdot (Rky \cdot Py^T)^*) \cdot Qky^*
\end{aligned}
$$

Now multiply $(Rkx \cdot Px^T) \cdot (Rky \cdot Py^T)^*$ in the following way:

$B = (Rkx \cdot Px^T) \cdot (Rky \cdot Py^T)^*$ ($\mathcal{O}(mk^2)$ flops)

Now take the svd of $B \in \mathbb{C}^{k \times k}$ ($\mathcal{O}(k^3)$ flops)

[Ub, S, Vb] = svd(B)

$$
\begin{aligned}
\% \quad A &= Qkx \cdot (Ub \cdot S \cdot Vb^*) \cdot Qky^* \\
\% \quad &= (Qkx \cdot Ub) \cdot S \cdot (Vb^* \cdot Qky^*) = (Qkx \cdot Ub) \cdot S \cdot (Qky \cdot Vb)^*
\end{aligned}
$$

Now multiply $Qkx \cdot Ub$ and $Qky \cdot Vb$ ($\mathcal{O}(mk^2)$ flops)

$U = Qkx \cdot Ub \quad , \quad V = Qky \cdot Vb$

Now return U, S, V, which is the skinny svd of A

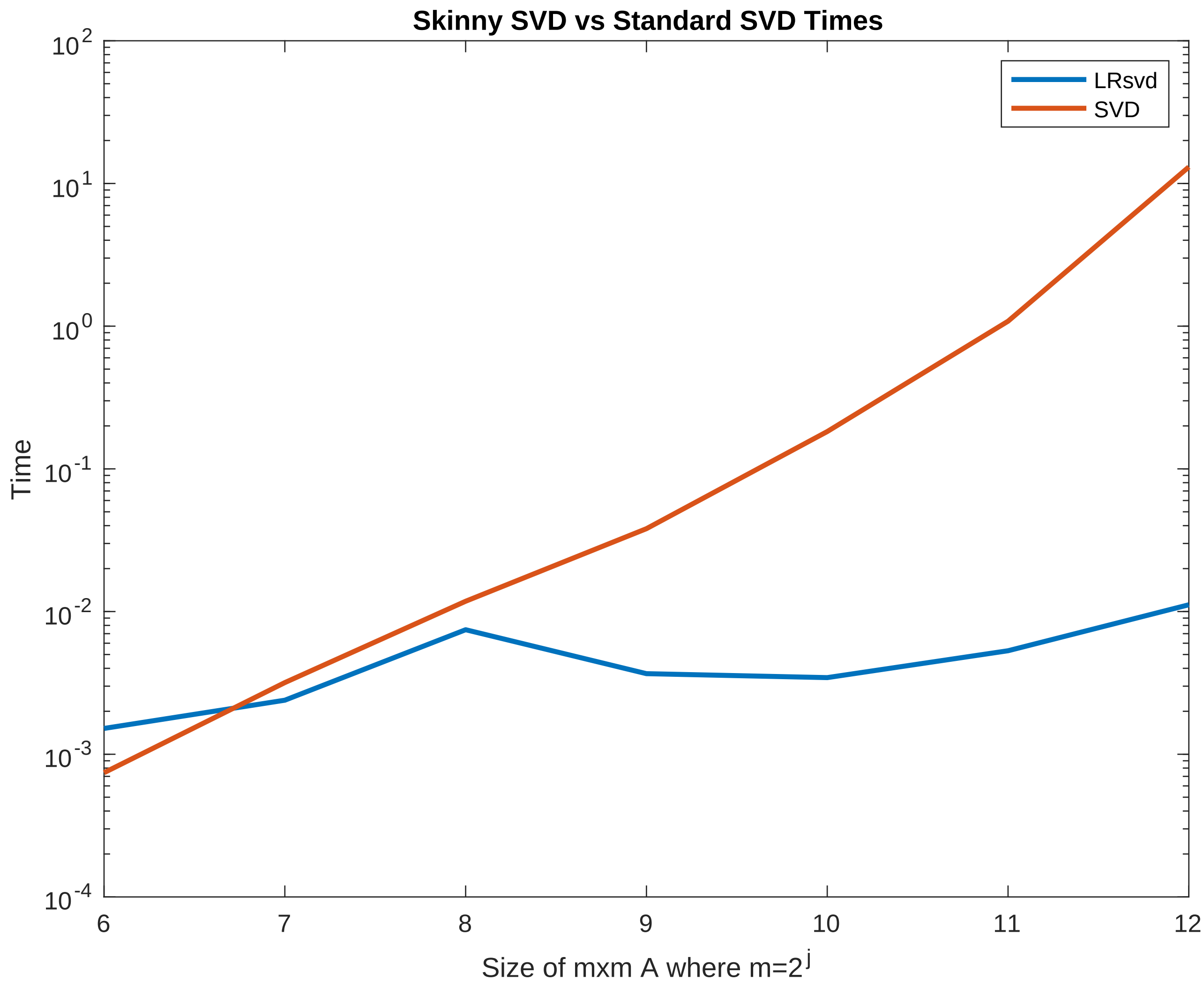
The total flops is on the order of:

$$\mathcal{O}(mk^2) + \mathcal{O}(mk^2) + \mathcal{O}(k^3) + \mathcal{O}(mk^2) = \mathcal{O}(mk^2 + k^3) \text{ flops}$$

```
>> testLRsvd

Error =

   1.1028e-14
```


Skinny SVD vs Standard SVD Times
LRsvd
SVD
Time
10^{2}
10^{1}
10^{0}
10^{-1}
10^{-2}
10^{-3}
10^{-4}
6
7
8
9
10
11
12
Size of mxm A where $m=2^j$