**Name:** Rajat Jaiswal **Entry Number:** 2017CS50415

1. Suppose $\mathbf{A}, \mathbf{B} \in \mathbb{C}^{m \times m}$ are two square matrices, and $\mathbf{C} = \mathbf{AB}$. Let the singular values of the three matrices be $a_1 \geq \cdots \geq a_m$, $b_1 \geq \cdots \geq b_m$, and $c_1 \geq \cdots \geq c_m$ respectively. Prove that $a_1 b_1 \geq c_1 \geq \max(a_1 b_m, a_m b_1)$, and similarly, $\min(a_1 b_m, a_m b_1) \geq c_m \geq a_m b_m$.

**Solution:** For, $A = U_A \Sigma_A V_A^*$ We know that $\|A\|_2 = max_x \dfrac{\|Ax\|_2}{\|x\|_2} = \|\Sigma_A\|_2 = max_i\{a_i\} = a_1$. This comes from Theorem 3.1 of Trefethen and Bau. And is also illustrated in Theorem 5.3. It is because 2-norm is unitarily invariant and essentially it becomes the norm of the diagonal matrix in the SVD. Similarly, $min_x \dfrac{\|Ax\|_2}{\|x\|_2} = max_i\{a_i\} = a_m$. Using submultiplicativity of induced p-norms,

$$\begin{gathered}\|C\|_2 = \|AB\|_2 \leq \|A\|_2 \cdot \|B\|_2 \implies \|C\|_2 \leq \|A\|_2 \cdot \|B\|_2 \\ \implies c_1 \leq a_1 \cdot b_1, \qquad \because \|A\|_2 = a_1, \|B\|_2 = b_1, \|C\|_2 = c_1\end{gathered}$$

We know that, $\max_x\{f(x) \cdot g(x)\} \geq max_x\{f(x)\} \cdot min_x\{g(x)\} \because g(x) \geq min_x\{g(x)\}$. $Bx$ is just a linear transformation of $x$, and is a vector.

$$\begin{gathered}\|C\|_2 = \max_x \frac{\|ABx\|_2}{\|x\|_2} = \max_x \frac{\|ABx\|_2 \cdot \|Bx\|_2}{\|Bx\|_2 \cdot \|x\|_2} \geq max_x \left\{\frac{\|ABx\|_2}{\|Bx\|_2}\right\} \cdot min_x \left\{\frac{\|Bx\|_2}{\|x\|_2}\right\} \\ \implies c_1 \geq a_1 \cdot b_m, \qquad \because \|C\|_2 = c_1, max_x \left\{\frac{\|ABx\|_2}{\|Bx\|_2}\right\} = a_1, min_x \left\{\frac{\|Bx\|_2}{\|x\|_2}\right\} = b_m\end{gathered}$$

Similarly,

$$\begin{gathered}\|C\|_2 = \max_x \frac{\|ABx\|_2}{\|x\|_2} = \max_x \frac{\|ABx\|_2 \cdot \|Bx\|_2}{\|Bx\|_2 \cdot \|x\|_2} \geq max_x \left\{\frac{\|Bx\|_2}{\|x\|_2}\right\} \cdot min_x \left\{\frac{\|ABx\|_2}{\|Bx\|_2}\right\} \\ \implies c_1 \geq b_1 \cdot a_m, \qquad \because \|C\|_2 = c_1, min_x \left\{\frac{\|ABx\|_2}{\|Bx\|_2}\right\} = a_m, max_x \left\{\frac{\|Bx\|_2}{\|x\|_2}\right\} = b_1.\end{gathered}$$

These two gives us the relation, $c_1 \geq max(a_1 b_m, b_1 a_m)$ and hence, $a_1 b_1 \geq c_1 \geq max(a_1 b_m, b_1 a_m)$.

Similarly, we can also say that $\min_x\{f(x) \cdot g(x)\} \geq min_x\{f(x)\} \cdot min_x\{g(x)\}$.

$$\begin{gathered}\min_x \frac{\|ABx\|_2}{\|x\|_2} = \min_x \frac{\|ABx\|_2 \cdot \|Bx\|_2}{\|Bx\|_2 \cdot \|x\|_2} \geq min_x \left\{\frac{\|ABx\|_2}{\|Bx\|_2}\right\} \cdot min_x \left\{\frac{\|Bx\|_2}{\|x\|_2}\right\} \\ \implies c_m \geq a_m \cdot b_m, \qquad \because \min_x \frac{\|ABx\|_2}{\|x\|_2} = c_m, min_x \left\{\frac{\|ABx\|_2}{\|Bx\|_2}\right\} = a_m, min_x \left\{\frac{\|Bx\|_2}{\|x\|_2}\right\} = b_m\end{gathered}$$

Similary, it can be said that $\min_x\{f(x) \cdot g(x)\} \leq min_x\{f(x)\} \cdot max_x\{g(x)\} \because g(x) \leq \max_x\{g(x)\}$.

$$\begin{gathered}\min_x \frac{\|ABx\|_2}{\|x\|_2} = \min_x \frac{\|ABx\|_2 \cdot \|Bx\|_2}{\|Bx\|_2 \cdot \|x\|_2} \leq min_x \left\{\frac{\|ABx\|_2}{\|Bx\|_2}\right\} \cdot max_x \left\{\frac{\|Bx\|_2}{\|x\|_2}\right\} \\ \implies c_m \leq a_m \cdot b_1, \qquad \because \min_x \frac{\|ABx\|_2}{\|x\|_2} = c_m, min_x \left\{\frac{\|ABx\|_2}{\|Bx\|_2}\right\} = a_m, max_x \left\{\frac{\|Bx\|_2}{\|x\|_2}\right\} = b_1 \\ \min_x \frac{\|ABx\|_2}{\|x\|_2} = \min_x \frac{\|ABx\|_2 \cdot \|Bx\|_2}{\|Bx\|_2 \cdot \|x\|_2} \leq max_x \left\{\frac{\|ABx\|_2}{\|Bx\|_2}\right\} \cdot min_x \left\{\frac{\|Bx\|_2}{\|x\|_2}\right\}\end{gathered}$$

*Continued on next page.*

$$\implies c_m \le a_1 \cdot b_m, \qquad \because \min_x \frac{\|ABx\|_2}{\|x\|_2} = c_m, max_x\left\{\frac{\|ABx\|_2}{\|Bx\|_2}\right\} = a_1, min_x\left\{\frac{\|Bx\|_2}{\|x\|_2}\right\} = b_m$$

These two gives us the relation, $c_m \le min(a_1 b_m, b_1 a_m)$ and hence, $a_m b_m \le c_m \le min(a_1 b_m, b_1 a_m)$.

2. Let $S$ be a subspace of $\mathbb{C}^m$. In the lectures, we defined its orthogonal complement as a subspace $T$ such that $S \cap T = \{0\}$, $S + T = \mathbb{C}^m$, and $S \perp T$. Show that this is precisely the set $\{\mathbf{v} \in \mathbb{C}^m : \mathbf{u}^*\mathbf{v} = 0\, \forall \mathbf{u} \in S\}$

**Solution:** We are given $S$, a subspace and $T$, its orthogonal complement.
Define $T' = \{\mathbf{v} \in \mathbb{C}^m : \mathbf{u}^*\mathbf{v} = 0\, \forall \mathbf{u} \in S\}$. We need to show that $T = T'$.

- Let $x \in T$ and by definition $x \in \mathbb{C}^m$. We know that, $\mathbf{u}^* x = 0, \forall \mathbf{u} \in S, \because T \perp S$ & $x \in T$.
  $\therefore$ by definition of $T'$, $x \in T'$. $\implies T \subseteq T'$.
- Let $x \in T'$ and $\because S + T = \mathbb{C}^m$ and are complement to each other, $\exists s_0 \in S, \exists t_0 \in T$, which are projection of $x$ in $S$ and $T$ respectively, such that $x = s_0 + t_0$. $\because \mathbf{u}^* x = 0,\ \forall \mathbf{u} \in S$
  $\implies \mathbf{u}^* s_0 + \mathbf{u}^* t_0 = 0, \because t_0 \in T, \mathbf{u}^* t_0 = 0 \implies \mathbf{u}^* s_0 = 0$. Since this is true $\forall \mathbf{u} \in S$, $\implies s_0 = 0$.
  $\implies x = t_0 \in T$, hence $x \in T \implies T' \subseteq T$.

Hence $T = T'$.

3. Consider a linearly independent set of $n$ real vectors $x_1, \cdots, x_n \in \mathbb{R}^m$. Suppose another set of vectors $y_1, \cdots, y_n \in \mathbb{R}^m$ is "congruent" to it, in the sense that all lengths and distances are equal: $\|x_i\|_2 = \|y_i\|_2$ for all $i$, and $\|x_i - x_j\|_2 = \|y_i - y_j\|_2$ for all $i \ne j$. Define the matrices $\mathbf{X} = [x_1, \cdots, x_n]$, and $\mathbf{Y} = [y_1, \cdots, y_n]$.
   (a) Prove that the reduced QR factorizations of $\mathbf{X}$ and $\mathbf{Y}$ have the same $\hat{\mathbf{R}}$.

**Solution:** Given, $\|x_i - x_j\|_2 = \|y_i - y_j\|_2$. Also, $\|x_i - x_j\|_2^2 = (x_i - x_j)^T \cdot (x_i - x_j)$. $\because$ All the vectors are real valued. This gives us that,

$$\|x_i - x_j\|_2^2 = (x_i^T - x_j^T) \cdot (x_i - x_j),\ \|y_i - y_j\|_2^2 = (y_i^T - y_j^T) \cdot (y_i - y_j)$$
$$\implies \|x_i - x_j\|_2^2 = x_i^T \cdot x_i + x_j^T \cdot x_j - 2 \cdot (x_i^T \cdot x_j),\ \implies \|y_i - y_j\|_2^2 = y_i^T \cdot y_i + y_j^T \cdot y_j - 2 \cdot (y_i^T \cdot y_j),$$
$$\because x_i^T \cdot x_j = x_j^T \cdot x_i$$

Given, $\|x_i - x_j\|_2^2 = \|y_i - y_j\|_2^2$, and $x_i^T \cdot x_i = \|x_i\|_2^2$, and $\|x_i\|_2^2 = \|y_i\|_2^2$, we get,

$$\implies \|x_i\|_2^2 + \|x_j\|_2^2 - 2 \cdot (x_i^T \cdot x_j) = \|y_i\|_2^2 + \|y_j\|_2^2 - 2 \cdot (y_i^T \cdot y_j)$$
$$\implies x_i^T \cdot x_j = y_i^T \cdot y_j$$
$$\implies \langle x_i, x_j \rangle = \langle y_i, y_j \rangle \text{ for all } i, j.$$

$\langle ., . \rangle$ denotes inner product. $\langle x_i, x_i \rangle = \langle y_i, y_i \rangle$ is already given to be true in the question.
Let $X' = X^T X$ and $Y' = Y^T Y$. $\because x_i$ denote $i^{th}$ column of $X$, $X'_{ij} = x_i^T \cdot x_j = \langle x_i, x_j \rangle$
Similarly, $Y'_{ij} = y_i^T \cdot y_j = \langle y_i, y_j \rangle$, using $\langle x_i, x_j \rangle = \langle y_i, y_j \rangle$, we get, $X'_{ij} = Y'_{ij} \implies X' = Y'$.

Let reduced QR factorization of $X = Q_x R_x$, and of $Y = Q_y R_y$, where $Q_x, Q_y$ are orthonormal matrices(column wise).

$$X' = Y' \implies X^T X = Y^T Y$$
$$\implies R_x^T Q_x^T Q_x R_x = R_y^T Q_y^T Q_y R_y$$
$$\implies R_x^T R_x = R_y^T R_y, \because Q_y^T Q_y = Q_x^T Q_x = I.$$
$$\implies R_x = R_y. \because R_x, R_y \text{ are upper triangular matrices.}$$

This can be easily shown for upper triangular matrices by multiplying and comparing element wise. $r_{x11}^2 = r_{y11}^2 \implies r_{x11} = r_{y11}$. Similarly, $r_{x11} \cdot r_{x12} = r_{y11} \cdot r_{y12} \implies r_{x12} = r_{y12}$, and so on.

(b) Give an algorithm to find an orthogonal matrix $\mathbf{Q}$ such that $\mathbf{Q}x_i = y_i$ for all $i$.

**Solution:** Let full QR factorization of $X = Q_{fx}R_{fx}$, and of $Y = Q_{fy}R_{fy}$, where $Q_{fx}, Q_{fy}$ are orthonormal square matrices(column wise) of size $m \times m$. This means that $Q_{fx}^T Q_{fx} = Q_{fx}Q_{fx}^T = I$. From part (a) we know that, $R_x = R_y$. In full QR factorization, $R_{fx}$ is constructed from $R_x$ by extending the last $m-n$ rows of $R_x$ with zeros so that $R_x$ extends to $R_{fx}$ of size $m \times n$, and similarly $R_{fy}$ is constructed from $R_y$, hence $R_{fx} = R_{fy}$.

Given that $\mathbf{Q}x_i = y_i, \ \forall i$. So we can easily extend this and see that $\mathbf{Q}[x_1, x_2, \cdots, x_n] = [y_1, y_2, \cdots, y_n] \implies \mathbf{QX} = \mathbf{Y}$.

$$\begin{aligned} &\mathbf{QX} = \mathbf{Y} \\ \implies &\mathbf{Q}Q_{fx}R_{fx} = Q_{fy}R_{fy} \\ \implies &\mathbf{Q}Q_{fx} = Q_{fy} \quad \because R_{fx} = R_{fy} \text{ and are upper triangular matrices} \\ \implies &\mathbf{Q} = Q_{fy}Q_{fx}^T \qquad \because Q_{fx}Q_{fx}^T = I \end{aligned}$$

**Algorithm 1** Solving for $\mathbf{QX} = \mathbf{Y}$

```
function GetQ(X, Y)
    Q_fx, R_fx ← fullQR(X)
    Q_fy, R_fy ← fullQR(Y)
    Q ← Q_fy · Q_fx^T
    return Q
end function
```

- **Running time analysis:** Full QR factorization using Gram-Schmidt takes $O(mn^2)$ time. Taking the transpose and matrix multiplication takes $O(m^3)$ time. Hence the overall algorithm takes $O(m(m^2 + n^2))$ time. $\because m \geq n$, we can say the overall running time is $O(m^3)$.

4. Consider a matrix $\mathbf{A} \in \mathbb{C}^{m \times n}$ and a vector $v \in \mathbb{C}^m$. Let $\mathbf{F} = \mathbf{I} - 2\dfrac{vv^*}{v^*v}$. .

(a) Show that $\mathbf{FA} = \mathbf{A} + vw^*$ for some vector $w$. Find the asymptotic operation count for both ways of computing $\mathbf{FA}$: (i) first computing $\mathbf{F}$ and then performing matrix multiplication, vs. (ii) first computing $w$, then $vw^*$, and then matrix addition.

**Solution:** Given $\mathbf{F} = \mathbf{I} - 2\dfrac{vv^*}{v^*v} \implies \mathbf{FA} = \mathbf{A} - 2\dfrac{vv^*\mathbf{A}}{v^*v} = \mathbf{A} + v\dfrac{-2v^*\mathbf{A}}{v^*v} \implies w^* = \dfrac{-2v^*\mathbf{A}}{v^*v}$.
Hence such a vector $w$ exists, and $w = \dfrac{-2\mathbf{A}^*v}{v^*v}$

(i) Computing $v^*v$ takes $O(m)$ flops and computing $vv^*$ takes $O(m^2)$ flops and division of $vv^*$ by $v^*v$ and then subtraction it from $\mathbf{I}$ takes $O(m^2)$ flops, hence computing $\mathbf{F}$ takes $O(m^2)$ flops. Multplication of $\mathbf{F}$ with $\mathbf{A}$ means multiply $m \times m$ matrix with $m \times n$ matrix which takes $O(m^2 n)$ flops, hence overall this way takes $O(m^2 n)$ flops.

(ii) For computing $w$, we need to compute $v^*v$ which takes $O(m)$ flops, and $\mathbf{A}^*v$ which takes $O(nm)$ flops. Computing $vw^*$ will take $O(mn)$ flops and then adding $vw^*$ to $\mathbf{A}$ also takes $O(mn)$ flops, hence this way overall computation of $\mathbf{FA}$ takes $O(mn)$ flops.

(b) Suppose we use an approximate vector $\tilde{v}$ and obtain $\tilde{\mathbf{F}} = \mathbf{I} - 2\dfrac{\tilde{v}\tilde{v}^*}{\tilde{v}^*\tilde{v}}$ instead. Show that if $\dfrac{\|\tilde{v} - v\|_2}{\|v\|_2} = O(\epsilon_m)$, then $\|\tilde{\mathbf{F}} - \mathbf{F}\|_2 = O(\epsilon_m)$, and $fl(\tilde{\mathbf{F}}\mathbf{A}) = \mathbf{F}(\mathbf{A} + \delta\mathbf{A})$ for some $\delta\mathbb{A}$ with $\dfrac{\|\delta A\|_2}{\|A\|_2} = O(\epsilon_m)$.

**Solution:** Take $\tilde{v} = v(1+\epsilon_1), \epsilon_1 \leq \epsilon_m$. Let's say inner product of two vectors produce an error of $\epsilon_3 \leq \epsilon_m$, and subtraction of two matrices produces an error of $\epsilon_4 \leq \epsilon_m$.
$\implies \tilde{\mathbf{F}} = \left(\mathbf{I} - 2 \cdot \dfrac{vv^*(1+\epsilon_1)^2(1+\epsilon_3)}{v^*v(1+\epsilon_1)^2(1+\epsilon_3)}\right)(1+\epsilon_4)$. Also, $(1+\epsilon)^{-1} = (1+\epsilon^{'})$ such that $\epsilon, \epsilon^{'} \leq \epsilon_m$, and we will ignore all terms of $O(\epsilon_m^2)$.

$$\implies \|\tilde{\mathbf{F}} - \mathbf{F}\|_2 = \left\| \mathbf{I}\epsilon_4 - 2 \cdot \frac{vv^*}{v^*v}\left((1 + 2\epsilon_1 + \epsilon_3 + \epsilon_4 + 2\epsilon_1^{'} + \epsilon_3^{'}) - 1\right) \right\|_2$$

Now, We will use submultiplicativity of induced 2-norm, and $\|v^*\|_2 = \|v\|_2$ and $v^*v = \|v\|_2^2$

$$\implies \|\tilde{\mathbf{F}} - \mathbf{F}\|_2 \leq \|\mathbf{I}\epsilon_\mathbf{4}\|_2 + 2\left(2\epsilon_1 + \epsilon_3 + \epsilon_4 + 2\epsilon_1^{'} + \epsilon_3^{'}\right) \cdot \left\| \frac{vv^*}{v^*v} \right\|_2$$

$$\implies \|\tilde{\mathbf{F}} - \mathbf{F}\|_2 \leq \epsilon_4 + \frac{2\left(2\epsilon_1 + \epsilon_3 + \epsilon_4 + 2\epsilon_1^{'} + \epsilon_3^{'}\right)}{\|v\|_2^2} \cdot \|vv^*\|_2$$

$$\implies \|\tilde{\mathbf{F}} - \mathbf{F}\|_2 \leq \epsilon_4 + \frac{2\left(2\epsilon_1 + \epsilon_3 + \epsilon_4 + 2\epsilon_1^{'} + \epsilon_3^{'}\right)}{\|v\|_2^2} \cdot \|v\|_2 \cdot \|v^*\|_2$$

$$\implies \|\tilde{\mathbf{F}} - \mathbf{F}\|_2 \leq \epsilon_4 + \frac{2\left(2\epsilon_1 + \epsilon_3 + \epsilon_4 + 2\epsilon_1^{'} + \epsilon_3^{'}\right)}{\|v\|_2^2} \cdot \|v\|_2^2$$

$$\implies \|\tilde{\mathbf{F}} - \mathbf{F}\|_2 \leq 15 \cdot \epsilon_m \implies \|\tilde{\mathbf{F}} - \mathbf{F}\|_2 = O(\epsilon_m)$$

---

Let's call $\beta = (2\epsilon_1 + \epsilon_3 + \epsilon_4 + 2\epsilon_1^{'} + \epsilon_3^{'}) = O(\epsilon_m)$. $fl(.)$ introduces error of $\epsilon_5 \leq \epsilon_m$
$\tilde{\mathbf{F}} = \mathbf{I}(1+\epsilon_4) - 2(1+\beta)\dfrac{vv^*}{v^*v}$.
$\mathbf{F^2} = \left(\mathbf{I} - 2\dfrac{vv^*}{v^*v}\right)\left(\mathbf{I} - 2\dfrac{vv^*}{v^*v}\right) = \mathbf{I} + 4\dfrac{\|v\|_2^2 vv^*}{\|v\|_2^4} - 4\dfrac{vv^*}{\|v\|_2^2} = \mathbf{I}$. $\|\mathbf{F}\|_\mathbf{2} \leq \|\mathbf{I}\|_2 + 2\dfrac{\|v\|_2 \cdot \|v^*\|_2}{\|v\|_2^2} = 3$.

$$\implies \tilde{\mathbf{F}}\mathbf{A} = \mathbf{A}(1+\epsilon_4) - 2(\mathbf{A} + \mathbf{A}\beta)\frac{vv^*}{v^*v}.$$

$$\implies fl(\tilde{\mathbf{F}}\mathbf{A}) = \mathbf{A}(1+\epsilon_4)(1+\epsilon_5) - 2(\mathbf{A} + \mathbf{A}\beta)(1+\epsilon_5)\frac{vv^*}{v^*v}.$$

Let $(1+\epsilon_4)(1+\epsilon_5) = 1 + \epsilon_6, \epsilon_6 = O(\epsilon_m)$ and $(1+\beta)(1+\epsilon_5) = 1 + \epsilon_7, \epsilon_7 = O(\epsilon_m)$

$$\implies fl(\tilde{\mathbf{F}}\mathbf{A}) = \mathbf{F}\mathbf{A} + \mathbf{A}\left(\epsilon_6 - 2\epsilon_7\frac{vv^*}{v^*v}\right).$$

$$\implies fl(\tilde{\mathbf{F}}\mathbf{A}) = \mathbf{F}\left(\mathbf{A} + \mathbf{F}\mathbf{A}\left(\epsilon_6 - 2\epsilon_7\frac{vv^*}{v^*v}\right)\right) \qquad \because \mathbf{F^2} = \mathbf{I}.$$

$$\implies \delta\mathbf{A} = \mathbf{F}\mathbf{A}\left(\epsilon_6 - 2\epsilon_7\frac{vv^*}{v^*v}\right).$$

$$\implies \|\delta\mathbf{A}\|_2 \leq \left(\epsilon_6 + 2\epsilon_7\frac{\|v\|_2 \cdot \|v^*\|_2}{\|v\|_2^2}\right)\|\mathbf{F}\|_2\|\mathbf{A}\|_2.$$

$$\implies \|\delta\mathbf{A}\|_2 \leq (\epsilon_6 + 2\epsilon_7) \cdot 3 \cdot \|\mathbf{A}\|_2.$$

$$\implies \frac{\|\delta\mathbf{A}\|_2}{\|\mathbf{A}\|_2} \leq 3 \cdot (\epsilon_6 + 2\epsilon_7).$$

$$\implies \frac{\|\delta\mathbf{A}\|_2}{\|\mathbf{A}\|_2} = O(\epsilon_m).$$

5. Suppose we want to find the general least-squares solution to the linear system $\mathbf{Ax} = \mathbf{b}$, where $\mathbf{A} \in \mathbb{C}^{m\times n}$ has $m > n$ and $rank(A) = r < n$. Let the full SVD of $\mathbf{A}$ be $\mathbf{A} = \mathbf{U\Sigma V}^*$.

   (a) Give an explicit formula for the unique vector $y \in range(\mathbf{A})$ which minimizes $\|b - y\|_2$.

**Solution:** We want to minimize $\|b - y\|_2$, so we need a point $y \in range(\mathbf{A})$ to $b$, so that norm of $r = b - y$ is minimized. It is clear geometrically that this will occur when $y = Pb$, where $P \in \mathbb{C}^{m\times m}$ that maps $\mathbb{C}^m$ onto $range(\mathbf{A})$. To minimize r, $r \perp range(\mathbf{A})$. This has also been claimed in Trefethan and Bau(Chapter 11) and can be seen in the figure below.

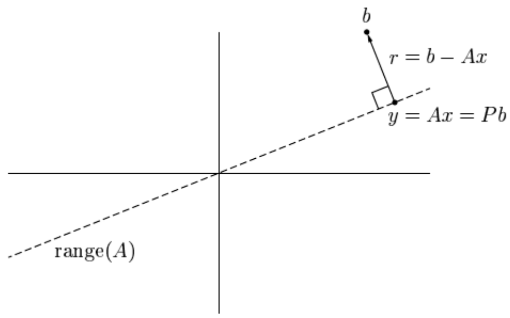


Figure 1: Minimizing $\|b - y\|_2$ in terms of orthogonal projections (Source: Trefethan and Bau).

Therefore, $y = Pb$. Now we need to construct $P$. We are given the full SVD of $\mathbf{A} = \mathbf{U\Sigma V}^*$. But it is also given that A has rank $r < n$. As proved in chapter 6 of Trefethan and Bau, the projector matrix $P$ for $range(\mathbf{A})$ can be obtained by $P = \hat{\mathbf{U}}\hat{\mathbf{U}}^*$, where $\hat{\mathbf{U}}$ comes from reduced SVD of $\mathbf{A} = \hat{\mathbf{U}}\hat{\mathbf{\Sigma}}\hat{\mathbf{V}}^*$.

We are given the full SVD of $\mathbf{A}$ and also given the rank $r$, we can construct $\hat{\mathbf{U}}$ from $\mathbf{U}$ by taking the first $r$ columns of $\mathbf{U}$ which is what we do in reduced SVD. Let $\mathbf{U_r}$ be the matrix $\mathbf{U}[:, :\mathbf{r}]$ which is the first $r$ columns of $\mathbf{U}$. Then by construction of reduced SVD $\hat{\mathbf{U}} = \mathbf{U_r}$. Hence $P = \mathbf{U_r}\mathbf{U_r^*}$.

$$\implies y = \mathbf{U_r}\mathbf{U_r^*}\mathbf{b}.$$

   (b) Find the general solution of the least-squares problem, i.e. all vectors $x$ for which $\|b - \mathbf{A}x\|_2$ is minimal. Which of these vectors minimizes $\|x\|_2$?

**Solution:** The approach for the solution has been taken form here. We are given $\mathbf{A} = \mathbf{U\Sigma V}^*$, where $\mathbf{U}$ and $\mathbf{V}$ are unitary square matrices of size $m \times m$ and $n \times m$ respectively. We know that 2-norm is unitarily invariant, $\implies \|b - \mathbf{A}x\|_2 = \|\mathbf{U}^*(b - \mathbf{A}x)\|_2$,. Define $\mathbf{Z} = \mathbf{V}^*x$.

$$\|b - \mathbf{A}x\|_2 = \|\mathbf{U}^*b - \mathbf{\Sigma V}^*x)\|_2, \because \mathbf{A} = \mathbf{U\Sigma V^*}, \mathbf{U^*U} = \mathbf{I}$$
$$\|b - \mathbf{A}x\|_2^2 = \Sigma_{i=1}^r (\sigma_i z_i - u_i^* b)^2 + \Sigma_{i=r+1}^n (u_i^* b)^2$$

To minimize $\|b - \mathbf{A}x\|_2^2$, we get, $\sigma_i z_i = u_i^* b,\ \forall i \in [1, 2, \cdots, r] \implies z_i = \dfrac{u_i^* b}{\sigma_i},\ \forall i \in [1, 2, \cdots, r]$ and $z_i =$ arbitrary $\forall i \in [r+1, r+2, \cdots, n]$. And $\min(\|b - \mathbf{A}x\|_2^2) = \Sigma_{i=r+1}^n (u_i^* b)^2$. We took, $\mathbf{Z} = \mathbf{V}^*x$, hence $x = \mathbf{VZ}$, where $\mathbf{Z}$ is as constructed. $x = [x_1, x_2, \cdots, x_n]$, where $x_i = \Sigma_{k=1}^n V_{ik} z_k$, $z_k$ as above. $x = \mathbf{VZ} \implies \|x\|_2 = \|\mathbf{VZ}\|_2 = \|\mathbf{Z}\|_2$. Since 2-norm is unitarily invariant. $\|\mathbf{Z}\|_2$ is minimum when $z_i = 0,\ \forall i \in [r+1, r+2, \cdots, n]$, and not arbitrary and hence $min(\|x\|_2) = \sqrt{\displaystyle\sum_{i=1}^r \left(\frac{u_i^* b}{\sigma_i}\right)^2}$.

In such a case for minimum $\|x\|_2$, $x = [x_1, x_2, \cdots, x_n]$, where $x_i = \displaystyle\sum_{k=1}^r V_{ik} z_k = \sum_{k=1}^r V_{ik} \left(\frac{u_k^* b}{\sigma_k}\right)$

6. Consider the set $P_n$ of all polynomials of degree $\leq n$ with complex coefficients. Any such polynomial $p$ can be represented as a coefficient vector $[p] \in \mathbb{C}^{n+1}$ via

$$p(x) = a_0 + a_1 x + \cdots + a_n x^n \iff [p] = \begin{bmatrix} a_0 \\ a_1 \\ \vdots \\ a_n \end{bmatrix}$$

Suppose we define an inner product on polynomials as $(p, q) = \int_{-1}^{1} \overline{p(x)} q(x)\, dx$. Thus two polynomials $p, q$ are orthogonal if $\int_{-1}^{1} \overline{p(x)} q(x)\, dx = 0$.

(a) Show that there exists a Hermitian matrix $\mathbf{G}$ such that $(p, q) = [p]^* \mathbf{G} [q]$ for all polynomials $p, q \in P_n$.

**Solution:** Proof by construction. Let $p(x) = \Sigma_{i=0}^{n} a_i x^i, q(x) = \Sigma_{i=0}^{n} b_i x^i$.

$$\overline{p(x)} q(x) = \Sigma_{i,j} a_i^* b_j x^{i+j}.$$

In case $i + j$ is odd, the term $a_i^* b_j$ won't contribute in integral as that term will become $0 (\because$ limits are -1 and 1), and when $i + j$ is even $a_i^* b_j$ will be multiplied by 2 and divided by $(i + j + 1)$.

So matrix $G$ of size $(n + 1) \times (n + 1)$ [indexing from 0] is, $G_{ij} = \begin{cases} 0 & \text{when i+j is odd} \\ \dfrac{2}{i + j + 1} & \text{when i+j is even} \end{cases}$

$G$ is hermitian because $G^* = G$, $\because G_{ij} = \dfrac{2}{i + j + 1} = \dfrac{2}{j + i + 1} = G_{ji}$ when $i + j$ is even, and when $i + j$ is odd $G_{ij} = G_{ji} = 0$, and since $G$ is real-valued, the complex conjugate remains the same. $G_{ij}$ is essentially multiplied by $i^{th}$ element of $[p]^*$ i.e. $a_i^*$ and $j^{th}$ element of $[q]$ i.e. $b_j$, and finally everything is summed over and hence $[p]^* \mathbf{G} [q] = \int_{-1}^{1} \overline{p(x)} q(x)\, dx = (p, q)$

(b) If $p, q \in P_n$ are two nonzero polynomials, what does it mean to project $q$ orthogonally onto the subspace $\langle p \rangle$ with respect to this inner product? Give an algebraic definition of orthogonal projection, and a formula for the corresponding matrix that acts on the coefficient vector $[q]$.

**Solution:** The subspace of p, $\langle p \rangle$ is defined as $\{k \cdot p | k \in \mathbb{C}\}$. When $q$ is projected orthogonally on $\langle p \rangle$ it means there is a component of $q$ on $\langle p \rangle$ say $(k' \cdot p)$ for some $k' \in \mathbb{C}$, and a component orthogonal to $\langle p \rangle$ which will be $(q - k' \cdot p)$. Since $(k' \cdot p)$ and $(q - k' \cdot p)$ are orthogonal to each other $\therefore (k' \cdot p, q - k' \cdot p) = 0$.

$$(k' \cdot p,\ q - k' \cdot p) = 0$$
$$\implies (k' \cdot p,\ q) - (k' \cdot p,\ k' \cdot p) = 0 \qquad \because [\overline{p(x)}(q(x) - r(x)) = \overline{p(x)} q(x) - \overline{p(x)} r(x)]$$
$$\implies \overline{k'} \cdot (p,\ q) = \overline{k'} k' \cdot (p,\ p) = 0 \qquad \because [\overline{k' \cdot p(x)} = \overline{k'} \cdot \overline{p(x)}\ \&\ \overline{p(x)}(k' \cdot q(x)) = k' \cdot \overline{p(x)} q(x)]$$
$$\implies k' = \frac{(p, q)}{(p, p)}$$

Hence the projection of $[q]$ on $\langle p \rangle$ is $\left( [p] \cdot \dfrac{(p, q)}{(p, p)} \right)$ and orthogonal projection is $\left( [q] - [p] \cdot \dfrac{(p, q)}{(p, p)} \right)$.

We know that, $(p, q) = [p]^* \mathbf{G} [q]$, using that we get, $[p] \cdot \dfrac{(p, q)}{(p, p)} = \dfrac{[p][p]^* \mathbf{G} [q]}{[p]^* \mathbf{G} [p]} = \dfrac{[p][p]^* \mathbf{G}}{[p]^* \mathbf{G} [p]} \cdot [q]$.

Hence the projection matrix for $\langle p \rangle$ acting on the coefficient vector $[q]$ is $\dfrac{[p][p]^* \mathbf{G}}{[p]^* \mathbf{G} [p]}$.

(c) Given a set of polynomials $p_1, \cdots, p_k$, we can now apply a Gram-Schmidt procedure to obtain a set of orthogonal polynomials (cf. Trefethen and Bau 7). Design and implement such an algorithm as a function $\mathbf{Q} = \texttt{orthogonalizePolynomials}(P)$, which takes a matrix $\mathbf{P}$ containing the coefficients of the polynomials $[p_j]$ as columns, and returns an analogous matrix $\mathbf{Q}$ for the orthogonal polynomials.

Apply your function to an identity matrix (representing the polynomials $(1, x, x^2, \cdots)$, and verify that the coefficients you obtain represent multiples of the Legendre polynomials.

**Solution:** Submitted `q6_2017CS50415.py` file. The function was tested on an identity matrix of size 5. The matrix Q obtained was as follows:

$$\begin{bmatrix} 0.70710678 & 0.00000000 & -0.79056942 & 0.00000000 & 0.79549513 \\ 0.00000000 & 1.22474487 & 0.00000000 & -2.80624304 & 0.00000000 \\ 0.00000000 & 0.00000000 & 2.37170825 & 0.00000000 & -7.95495129 \\ 0.00000000 & 0.00000000 & 0.00000000 & 4.67707173 & 0.00000000 \\ 0.00000000 & 0.00000000 & 0.00000000 & 0.00000000 & 9.28077650 \end{bmatrix}$$

As evident, the columns are the multiples of the Legendre polynomials. For e.g, Column 3 represents the polynomial $Q(x) = 2.37170825x^2 - 0.79056942$, which is equivalent to $1.58113883(\frac{3}{2}x^2 - \frac{1}{2}) = 1.58113883 \cdot P_3(x)$.