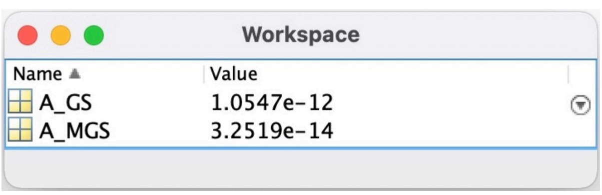
MATH 501 Homework 11

Qilin Ye, April 18, 2021

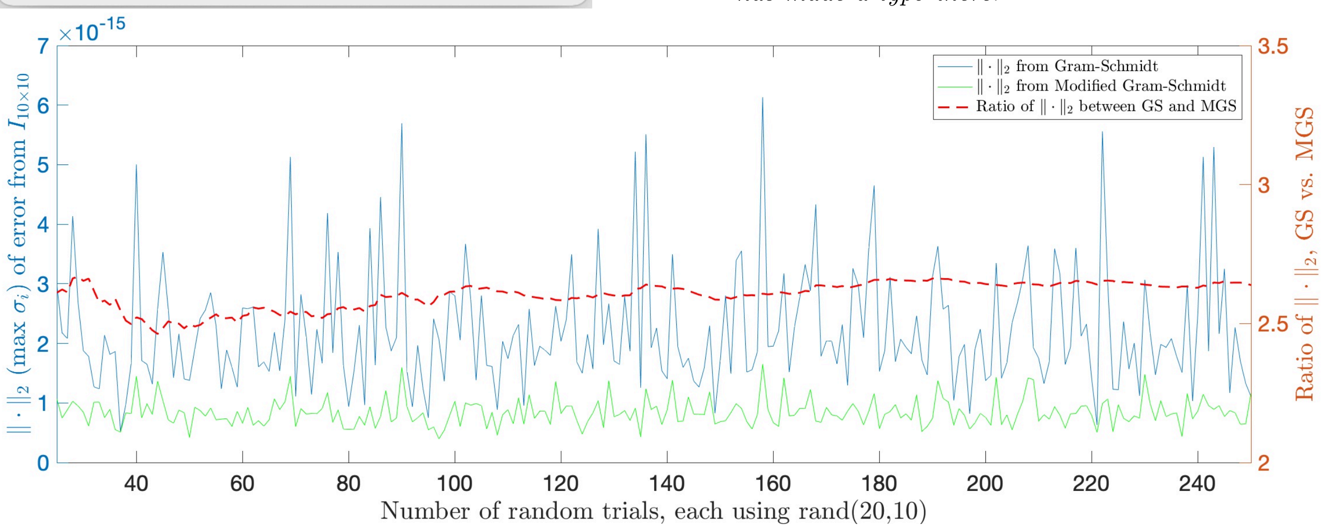
Pseudocode Implementation

```
A = zeros(20,10);
for i = 1:20
    for j = 1:10
        A(i,j) = ((2*i-21)/19)^(j-1);
    end
end
B = Gram_Schmidt(A);
C = Modified_Gram_Schmidt(A);
A_GS =norm(B.'*B - eye(10));
A_MGS =norm(C.'*C - eye(10));

u = zeros(50,1);
v = zeros(50,1);

for i = 1:250
    M = rand(20,10);
    M_GS = Gram_Schmidt(M);
    M_MGS = Modified_Gram_Schmidt(M);
    u(i) = norm(M_GS.' * M_GS - eye(10));
    v(i) = norm(M_MGS.' * M_MGS - eye(10));
end
(Some extra code to generate graphs)
```




```
function B = Gram_Schmidt(A)
    B = zeros(20,10);
    C = zeros(20,10);
    T = zeros(20,10);
    for j = 1:10
        for i = 1:j-1
            T(i,j) = dot(A(1:20,j),B(1:20,i));
        end
        C(1:20,j) = A(1:20,j);
        for i = 1:j-1
            C(1:20,j) = C(1:20,j) - T(i,j) * B(1:20,i);
        end
        T(j,j) = norm(C(1:20,j));
        B(1:20,j) = C(1:20,j) / T(j,j);
    end
end
function A = Modified_Gram_Schmidt(A)
    T = zeros(20,10);
    for k = 1:10
        A(1:20,k) = A(1:20, k) / norm(A(1:20,k));
        for j = k+1:10
            A(1:20,j) = A(1:20,j) - dot(A(1:20,k),
                A(1:20,j)) * A(1:20,k);
        end
    end
end
```

*The * differs from the book's $^{-1}$; I believe the book has made a typo there.*



Textbook Problems

5.3.3 Prove that if $A_{m\times n}$ is of rank n then A^*A is Hermitian and positive definite.

Proof. If A is of rank n then $Ax = 0$ if and only if $x = 0$. Thus for nonzero x,

$$x^*A^*Ax = (Ax)^*(Ax) = \|Ax\|^2 > 0.$$

That A^*A is Hermitian doesn't even require A to be rank n. It just follows directly from

$$(A^*A)^* = A^*A^{**} = A^*A.$$

□

5.3.6 Let $\{u_1, \dots, u_n\}$ be an orthonormal basis for a subspace U of an inner product space X. Define $P : X \to U$ by

$$P(x) = \sum_{i=1}^{n} \langle x, u_i \rangle u_i.$$

Prove that

(1) P is linear,

(2) P is idempotent,

(3) $Px = x$ if $x \in U$, and

(4) $\|Px\|_2 \leqslant \|x\|_2$ for all $x \in X$.

Proof. (1) Linearity directly follows from the fact that inner product itself is a linear mapping (and of course finite sums don't spoil the linearity).

(2) This follows from (3) as $P^2x = P(Px) = P(v)$ for some $v \in U$. So we'll only show (3).

(3) If $x \in U$ then $x = \sum_{i=1}^{n} c_i u_i$ for some $\{c_i\}_{i=1}^n$. Thus,

$$\begin{aligned} P(x) &= \sum_{i=1}^{n} \langle x, u_i \rangle u_i = \sum_{i=1}^{n} \left[\langle \sum_{j=1}^{n} c_j u_j, u_i \rangle u_i \right] \\ &= \sum_{i=1}^{n} \left[\sum_{j=1}^{n} \langle c_j u_j, u_i \rangle u_i \right] = \sum_{i=1}^{n} \sum_{j=1}^{n} c_j \langle u_j, u_i \rangle u_i \\ &= \sum_{i,j=1}^{n} c_j \delta_{i,j} u_i = \sum_{k=1}^{n} c_k u_k = x. \end{aligned}$$

(4) Since the question asks about $\|\cdot\|_2$ I shall assume that X can be identified with $\mathbb{R}^m$ for some m. Then we are able to extend $\{u_1, \dots, u_n\}$ to $u_1, \dots, u_n, \dots, u_m$ which forms an orthonormal basis for X. Then the claim becomes clear:

$$\|Px\|_2^2 = \left\| \sum_{i=1}^{n} \langle x, u_i \rangle u_i \right\|_2^2 = \sum_{i=1}^{n} \langle x, u_i \rangle^2 \leqslant \sum_{i=1}^{m} \langle x, u_i \rangle^2 = \|x\|_2^2.$$

□

5.3.17 Prove that if Q is unitary then for all x, y,

$$\|x\|_2 = \|Qx\|_2 \text{ and } \langle x, y \rangle = \langle Qx, Qy \rangle.$$

Also compute $\|Q\|_2$ using the matrix norm subordinate to the Euclidean norm.

Proof. For all x, $\|Qx\|_2^2 = \langle Qx, Qx\rangle = x^*Q^*Qx = x^*x = \|x\|_2$. To compute $\|Q\|_2$, on one hand we have

$$\|Qx\|_2 = \|x\|_2 \implies \|Q\|_2 \leqslant 1$$

and on the other hand, letting $\tilde{x}$:= any eigenvector of Q gives $\|Q\tilde{x}\| = \|\tilde{x}\| \implies \|Q\|_2 \geqslant 1$. Thus $\|Q\|_2 = 1$. □

5.3.19 Let A be an $m \times n$ matrix, b an m-vector, and $\alpha > 0$. Using the Euclidean norm, define

$$F(x) := \|Ax - b\|_2^2 + \alpha\|x\|_2^2.$$

Prove $F(x)$ is a minimum when x is a solution of the equation

$$(A^TA + \alpha I)x = A^Tb.$$

Prove that when x is so defined,

$$F(x+h) = F(x) + (Ah)^TAh + \alpha h^Th.$$

Proof. It suffices to prove the second claim directly, after which the first claim follows since

$$(Ah)^TAh + \alpha h^Th = \|Ah\|_2^2 + \alpha\|h\|_2^2 \geqslant 0.$$

Indeed,

$$\begin{aligned} F(x+h) &= \|A(x+h) - b\|_2^2 + \alpha\|x+h\|_2^2 \\ &= \|Ax-b\|_2^2 + \|Ah\|_2^2 + 2\langle Ax - b, Ah\rangle + \alpha\|x\|_2^2 + \alpha\|h\|_2^2 + 2\alpha\langle x, h\rangle \\ &= \|Ax-b\|_2^2 + \|Ah\|_2^2 + 2\left\langle A^T(Ax-b), h\right\rangle + \alpha\|x\|_2^2 + \alpha\|h\|_2^2 + 2\alpha\langle x, h\rangle \\ &= F(x) + (Ah)^TAh + \alpha h^Th + 2\underbrace{\left\langle A^T(Ax-b), h\right\rangle + 2\langle \alpha Ix, h\rangle}_{=0 \text{ since } A^T(Ax-b)+2\alpha Ix=0}. \end{aligned}$$

□

5.3.24 Show that in solving the least-squares problem for $Ax = b$, we can replace the normal equations by $CAx = Cb$ where C is any $n \times m$ matrix row-equivalent to A^T.

Proof. Indeed, we know that the least-squares solution to $Ax = b$ is the x that solves $A^T(Ax - b) = 0$. Since A^T and C are equivalent, $C = EA^T$ for some invertible invertible $n \times n$ matrix E. Hence x solves $A^T(Ax - b) = 0$ if and only if x solves $EA^T(Ax - b) = C(Ax - b) = 0$. □

5.3.25 Let A be an $m \times n$ matrix of rank n. Let b be any point in $\mathbb{R}^m$. Show that the sets

$$K_\lambda := \{x \in \mathbb{R}^n : \|Ax - b\|_2 \leqslant \lambda\}$$

are closed and bounded.

Proof. Let $\lambda \geqslant 0$ be given and fix it. We first show the closure of K_λ. Suppose $(x_1, x_2, \ldots) \subset K_\lambda$ converges to some $x \in \mathbb{R}^n$. By triangle inequality,

$$\|Ax - b\| \leqslant \|Ax - Ax_i\| + \|Ax_i - b\| \text{ for } x_i \in (x_1, x_2, \ldots).$$

The first term on the RHS converges to 0 because A is a bounded operator (linear with finite-dimensional domain) and the second term $\leqslant \lambda$. Therefore the sum $\leqslant \lambda$, i.e., $x \in K_\lambda$.

For boundedness, suppose K_λ is unbounded so there exists $(x_1, x_2, \ldots) \subset K_\lambda$ (most likely a different sequence from the one used previously...a bit of abuse of notation here) such that $\|x_k\| \geqslant k$. Triangle inequality gives

$\|Ax_k\| \leqslant \|Ax_k - b\| + \|b\|$. The RHS is bounded by some constant M in this case, so M must also bound the LHS. Thus $\|Ax_k\|$ is bounded for all $x_k \in (x_1, x_2, ...)$. It follows that if we define the sequence $(y_1, y_2, \ldots)$ by

$$y_k := \frac{x_k}{\|x_k\|}$$

then $Ay_k = Ax_k/\|x_k\|$ converges to 0 (the numerator is bounded and the denominator $\to \infty$). Since (I hope this is allowed in 501) unit balls in finite-dimensional spaces (in particular $\mathbb{R}^n$) are compact, $(y_1, y_2, ...)$ admits a convergent subsequence that converges to some y with $\|y\| = 1$. But then $Ay = 0$ for some nonzero y, contradicting the assumption that A is of full column rank. Thus K_λ is bounded. □

5.3.26 Assume the hypotheses in the preceding problem. Prove that if $\lambda = 2\|b\|_2$ then

$$\inf_{x\in\mathbb{R}^n} \|Ax - b\|_2 = \inf_{x\in K_\lambda} \|Ax - b\|_2.$$

Proof. That $\inf_{x\in\mathbb{R}^n} \|Ax - b\|_2 \leqslant \inf_{x\in K_\lambda} \|Ax - b\|_2$ is trivial, so it suffices to show that any $x \in \mathbb{R}^n \setminus K_\lambda$ has no effect on determining the infimum. Indeed, letting $x = 0$ tells us that the infimum of both sides are $\leqslant \|b\|_2$, so it suffices to check the infimum of all x's with $\|Ax - b\|_2 \leqslant \|b\|_2$ which, of course, is contained in K_λ. (Not sure why the problem asked explicitly for $2\|b\|_2$ though...) □

5.3.27 Show that if $A_{m\times n}$ is of rank n then the least-squares solution of $Ax = b$ satisfies the inequality

$$\|x\|_2 \leqslant 2\|b\|_2\|B\|_2$$

where B is any left inverse of A. Here $\|\cdot\|_2$ denotes the matrix norm subordinate to Euclidean $\|\cdot\|_2$.

Proof. If x solves $Ax = b$ then

$$\|x\|_2 = \|BAx\|_2 \leqslant \|B\|_2\|Ax\|_2 \leqslant 2\|B\|_2\|b\|_2.$$ □

5.3.28 Let A be an $m \times n$ matrix of unspecified rank. Let $b \in \mathbb{R}^m$ and let

$$\rho := \inf_{x\in\mathbb{R}^n} \|Ax - b\|.$$

Prove that this infimum is attained, regardless of the rank of A and the choice of $\|\cdot\|$.

Proof. If A is of full column rank, by (a slight generalization of) 5.3.26 $\inf_{x\in\mathbb{R}^n} \|Ax - b\|$ is the same as $\inf_{x\in K_\lambda} \|Ax - b\|$ where $\lambda = 2\|b\|$ (the proof of the generalized version, i.e., without specifying $\|\cdot\|_2$, is identical to that of the case $\|\cdot\|_2$). Since the map $x \mapsto \|Ax - b\|$ is the composition of several continuous maps (namely the composition of $\|\cdot\|$ with the sum of $x \mapsto Ax$ and the constant b), it is also continuous. A closed and bounded set K_λ in a finite-dimensional space is compact, so its image must be compact as well. Therefore the infimum is attained.

Now suppose A $(m \times n)$ is not of full column rank. Clearly if we swap the orders of the columns of A, the infimum is unaffected in either case (we can just swap the corresponding components of x as well). With that justified, assume all the pivot columns are aligned on the left side of A and suppose there are $k < n$ such columns. Now define $B_{n\times k}$ to be the left $n \times k$ sub-matrix of $I_{n\times n}$, i.e., the leftmost k columns of the $n \times n$ identity matrix. It follows that (1) A shares the same column space with AB and, more importantly,

(2) AB is of *full column rank*. Therefore there does exist $\tilde{x} \in \mathbb{R}^k$ such that

$$\inf_{x\in\mathbb{R}^k} \|ABx - b\| = \|AB\tilde{x} - b\|.$$

But we've said AB and A have the same column space, so $\inf_{x\in\mathbb{R}^k} \|ABx - b\| = \inf_{x\in\mathbb{R}^n} \|Ax - b\|$ and the infimum for the RHS can also be attained by $B\tilde{x} \in \mathbb{R}^n$. □

5.3.29 Adopt the assumptions in the preceding problem and prove that the equation $A^TAx = A^Tb$ has a solution, regardless of the rank of A.

Proof. This problem amounts to showing that A^Tb lies inside the column space of A^TA. Since the double orthogonal complement of a closed subspace is simple the subspace itself (i.e., if U is a closed subspace then $(U^\perp)^\perp = U$), it suffices to show that A^Tb is orthogonal to everything inside $(C(A^TA))^\perp$ (the orthogonal complement of the column space of A^TA). Let $y \in (C(A^TA))^\perp$ be arbitrarily chosen. Then

$$y^TA^TA = 0 \implies y^TA^TAy = \|Ay\|^2 = 0 \implies Ay = 0 \implies y^TA^Tb = 0,$$

i.e., y is indeed orthogonal to A^Tb. Thus $A^Tb \in C(A^TA)$, proving the claim. □

5.3.42 Determine $\kappa_\infty(A)$ and $\kappa_\infty(A^*A)$ where

$$A = \begin{bmatrix} 1 & 1 & 1 \\ \epsilon & 0 & 0 \\ 0 & \epsilon & 0 \\ 0 & 0 & \epsilon \end{bmatrix} \quad \text{and} \quad A^*A = \begin{bmatrix} 1+\epsilon^2 & 1 & 1 \\ 1 & 1+\epsilon^2 & 1 \\ 1 & 1 & 1+\epsilon^2 \end{bmatrix}.$$

What happens as $\epsilon \to 0$?

Solution Recall [or not] that $\kappa_\infty(A) = \|A\|_\infty \|A^+\|_\infty$. Putting A and A^* into *WolframAlpha*, we have

$$A^+ = \frac{1}{\epsilon^2+3} \begin{bmatrix} 1 & (\epsilon^2+2)/\epsilon & -1/\epsilon & -1/\epsilon \\ 1 & -1/\epsilon & (\epsilon^2+2)/\epsilon & -1/\epsilon \\ 1 & -1/\epsilon & -1/\epsilon & (\epsilon^2+2)/\epsilon \end{bmatrix}$$

and

$$(A^*A)^{-1} = \frac{1}{\epsilon^4+3\epsilon^2} \begin{bmatrix} \epsilon^2+2 & -1 & -1 \\ -1 & \epsilon^2+2 & -1 \\ -1 & -1 & \epsilon^2+2 \end{bmatrix}.$$

Recall that $\|M_{n\times n}\|_\infty = \max_{1\leqslant i\leqslant n} \sum_{j=1}^{n} |M_{i,j}|$ (the sum of entries of the row in which the sum of component-wise absolute values is the maximum). Now we compute $\kappa_\infty(A)$ and $\kappa_\infty(A^*A)$:

$$\kappa_\infty(A) = \|A\|_\infty \|A^+\|_\infty = 3 \cdot \frac{1+\epsilon}{3+\epsilon^2} = \frac{3+\epsilon}{3+\epsilon^2} \text{ and } \lim_{\epsilon\to 0} \kappa_\infty(A) = \lim_{\epsilon\to 0} \frac{3+\epsilon}{3+\epsilon^2} = 1,$$

and

$$\kappa_\infty(A^*A) = (3+\epsilon^2)\frac{\epsilon^2}{\epsilon^4+3\epsilon^2} = 1 \text{ and } \lim_{\epsilon\to 0} \kappa_\infty(A^*A) = 1.$$