August 21, 2009: I succeed in proving a thing I have believed: $\sqrt{\frac{2}{n+1}}[\sin\frac{\pi ij}{n+1}]$ *is unitary!*

Consider the Fourier matrix of order $2(n+1)$:

$$F_{2(n+1)} = \frac{1}{\sqrt{2(n+1)}}[\omega_{2(n+1)}^{ij}]_{i,j=0}^{2(n+1)-1}, \quad \omega_{2(n+1)} = e^{-\mathbf{i}\frac{2\pi}{2(n+1)}} = e^{-\mathbf{i}\frac{\pi}{n+1}}.$$

Note that, if $o_n = \sqrt{2/(n+1)}$, then

$$\begin{array}{l} F_{2(n+1)} = \frac{1}{2}(C - \mathbf{i}S), \\ c_{ij} = o_n \cos\frac{ij\pi}{n+1}, \quad s_{ij} = o_n \sin\frac{ij\pi}{n+1}, \quad i,j = 0,\dots,2(n+1)-1. \end{array}$$

Since S and C are real symmetric matrices, we have

$$I = F_{2(n+1)}^* F_{2(n+1)} = \frac{1}{2}(C+\mathbf{i}S)\frac{1}{2}(C-\mathbf{i}S) = \frac{1}{4}[(C^2+S^2) + \mathbf{i}(SC - CS)],$$

$$Q = F_{2(n+1)}^2 = \frac{1}{2}(C-\mathbf{i}S)\frac{1}{2}(C-\mathbf{i}S) == \frac{1}{4}[(C^2-S^2) - \mathbf{i}(CS+SC)],$$

being

$$Q = \begin{bmatrix} 1 & & & \\ & & & J \\ & & 1 & \\ & J & & \end{bmatrix}, \quad J \ n\times n \text{ counter-identity.}$$

As a consequence

$$\begin{array}{l} C^2 + S^2 = 4I \\ C^2 - S^2 = 4Q \\ CS = SC = 0 \end{array} \ \Rightarrow S^2 = 2(I-Q) = 2\begin{bmatrix} 0 & & & \\ & I & & -J \\ & & 0 & \\ & -J & & I \end{bmatrix}.$$

Now let S_{11}, S_{12}, S_{22} be the $n\times n$ matrices defined by the equality

$$S = \begin{bmatrix} 0 & & & \\ & S_{11} & & S_{12} \\ & & 0 & \\ & S_{12}^T & & S_{22} \end{bmatrix},$$

that is,

$$\begin{array}{c} (S_{11})_{rs} = o_n \sin\frac{rs\pi}{n+1}, \ (S_{12})_{rs} = o_n \sin\frac{r(n+1+s)\pi}{n+1}, \\ (S_{22})_{rs} = o_n \sin\frac{(n+1+r)(n+1+s)\pi}{n+1}, \ 1 \le r,s \le n. \end{array}$$

Observe that S_{11} and S_{22} are real symmetric and related by the identity $S_{22} = JS_{11}J$; moreover S_{12} is persymmetric, i.e. $S_{12}J = JS_{12}^T$. (Recall that S_{11} is the (sine) transform diagonalizing the algebra τ of all polynomials in

$$X = \begin{bmatrix} & 1 & & & \\ 1 & & 1 & & \\ & 1 & & \ddots & \\ & & \ddots & & 1 \\ & & & 1 & \end{bmatrix}$$

).

Since

$$S^2 = \begin{bmatrix} 0 & & & \\ & S_{11}^2 + S_{12}S_{12}^T & & S_{11}S_{12} + S_{12}S_{22} \\ & & 0 & \\ & S_{12}^T S_{11} + S_{22}S_{12}^T & & S_{12}^T S_{12} + S_{22}^2 \end{bmatrix},$$

we obtain four identities which in fact reduce to the following only two:

$$S_{11}^2 + S_{12}S_{12}^T = 2I, \quad S_{11}S_{12}J + S_{12}JS_{11} = -2I. \tag{1}$$

The sum of them yields $0 = S_{11}(S_{11}+S_{12}J)+S_{12}J(S_{11}+S_{12}J) = (S_{11}+S_{12}J)^2$, but this can happen only if

$$S_{11} + S_{12}J = 0, \quad S_{12} = -S_{11}J \tag{2}$$

(a real symmetric matrix with all the eigenvalues equal to 0 must be null).

Now we are near the thesis. In fact, by (2) the first identity in (1) becomes $2I = S_{11}^2 + (-S_{11}J)(-S_{11}J)^T = 2S_{11}^2$, and so $S_{11}^2 = I$.

Remark. From the equality $F_{2(n+1)} = \frac{1}{2}(C - \mathbf{i}S)$ it follows that $S = \mathbf{i}(F_{2(n+1)} - F^*_{2(n+1)}) = \mathbf{i}(I-Q)F_{2(n+1)}$. So, *the sine transform of* $\mathbf{z}$ $n \times 1$, $S_{11}\mathbf{z}$, *can be computed via a discrete Fourier transform of order* $2(n+1)$:

$$\mathbf{i}(I-Q)F_{2(n+1)} \begin{bmatrix} 0 \\ \mathbf{z} \\ 0 \\ \mathbf{0} \end{bmatrix} = \begin{bmatrix} 0 & & & \\ & S_{11} & & -S_{11}J \\ & & 0 & \\ & -JS_{11} & & JS_{11}J \end{bmatrix} \begin{bmatrix} 0 \\ \mathbf{z} \\ 0 \\ \mathbf{0} \end{bmatrix} = \begin{bmatrix} 0 \\ S_{11}\mathbf{z} \\ 0 \\ -JS_{11}\mathbf{z} \end{bmatrix}.$$

□ Investigate the four submatrices of C, perhaps they also can be expressed in terms of only one and this one is a transform diagonalizing some algebra of matrices ...

The matrix

$$A = \begin{bmatrix} 3 & 2 \\ 1 & 2 \end{bmatrix}$$

does not satisfy the equation $A^*A = AA^*$, thus there is no unitary matrix diagonalizing A. However, $T^{-1}AT$ is diagonal for a suitable T:

$$D^{-1}AD = \begin{bmatrix} 3 & \sqrt{2} \\ \sqrt{2} & 2 \end{bmatrix}, \quad D = \begin{bmatrix} \sqrt{2} & 0 \\ 0 & 1 \end{bmatrix},$$

$$\begin{bmatrix} \alpha & -\beta \\ \beta & \alpha \end{bmatrix} \begin{bmatrix} 3 & \sqrt{2} \\ \sqrt{2} & 2 \end{bmatrix} \begin{bmatrix} \alpha & \beta \\ -\beta & \alpha \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 4 \end{bmatrix}, \quad \alpha = \frac{1}{\sqrt{3}}, \ \beta = \sqrt{\frac{2}{3}},$$

$$T = \frac{1}{\sqrt{3}} \begin{bmatrix} \sqrt{2} & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & \sqrt{2} \\ -\sqrt{2} & 1 \end{bmatrix} = \frac{1}{\sqrt{3}} \begin{bmatrix} \sqrt{2} & 2 \\ -\sqrt{2} & 1 \end{bmatrix}.$$

The condition number of T (in the 2-norm), $\mu_2(T) = \|T\|_2\|T^{-1}\|_2$, is greater than 1:

$$T^*T = \frac{1}{3} \begin{bmatrix} 4 & \sqrt{2} \\ \sqrt{2} & 5 \end{bmatrix} \Rightarrow \|T\|_2 = \sqrt{\rho(T^*T)} = \sqrt{2},$$

$$T^{-1} = \frac{\sqrt{3}}{3\sqrt{2}} \begin{bmatrix} 1 & -2 \\ \sqrt{2} & \sqrt{2} \end{bmatrix},\ (T^{-1})^*(T^{-1}) = \begin{bmatrix} \frac{1}{2} & 0 \\ 0 & 1 \end{bmatrix} \Rightarrow \ \|T^{-1}\|_2 = 1.$$

So, $\mu_2(T) = \sqrt{2}$. Since $\|T\|_\infty = \frac{\sqrt{2}+2}{\sqrt{3}}$, $\|T^{-1}\|_\infty = \frac{\sqrt{3}}{\sqrt{2}}$, we have $\mu_\infty(T) = 1+\sqrt{2}$.

Can a non-unitary matrix T have condition number equal to 1 ?

If yes, then, by the Bauer-Fike theorem, the eigenvalue problem would be optimally conditioned for a class of matrices A larger than normal (the A diagonalized by T, $\mu_2(T) = 1$).

A $n \times n$ matrix A is said reducible if there exists $\mathcal{I} \subset \mathcal{N} = \{1, 2, \ldots, n\}$, $\mathcal{I} \neq \emptyset, N$, such that $a_{ik} = 0$ for all $i \in \mathcal{I}, k \in \mathcal{N}\backslash\mathcal{I}$. Equivalently, A is reducible if there exists a permutation matrix P such that

$$P^T AP = \begin{bmatrix} \square_{n-i} & * \\ 0 & \square_i \end{bmatrix},\ \square_k \ k \times k \ matrices,\ i \neq 0, n$$

$(i = |\mathcal{I}|, n - i = |\mathcal{N}\backslash\mathcal{I}|)$.

Set

$$C_i = \{z \in \mathbb{C} : |z - a_{ii}| < \sum_{j=1, j\neq i}^{n} |a_{ij}|\}.$$

It is well known that the subset $\cup_{i=1}^n \overline{C_i}$ of $\mathbb{C}$ includes all the eigenvalues of A (Gershgorin first theorem).

If A is not reducible then we can say something more:

If A is a irreducible $n\times n$ matrix and C_i are the inner parts of the Gershgorin disks, then the set $(\cup_{i=1}^n C_i) \cup (\cap_{i=1}^n \partial C_i)$ includes all the eigenvalues of A.

Proof. If λ is an eigenvalue of A, then $\sum_j a_{ij}x_j = \lambda x_i$, $\sum_{j,j\neq i} a_{ij}x_j = (\lambda - a_{ii})x_i$,

$$|\lambda - a_{ii}||x_i| \leq \sum_{j,j\neq i} |a_{ij}||x_j|, \quad \forall\, i.$$

Set $\mathcal{I} = \{j : |x_j| = \|\mathbf{x}\|_\infty\}$. Assume $\mathcal{I} \neq N$ and let $i \in \mathcal{I}$, $k \in \mathcal{N}\backslash\mathcal{I}$ such that $a_{ik} \neq 0$. Then

$$\begin{array}{rcl} |\lambda - a_{ii}||x_i| & \leq & \sum_{j,j\neq i} |a_{ij}||x_j| \\ & = & \sum_{j\in\mathcal{I},j\neq i} |a_{ij}||x_j| + |a_{ik}||x_k| + \sum_{j\in\mathcal{N}\backslash\mathcal{I},j\neq k} |a_{ij}||x_j| \\ & < & \sum_{j\in\mathcal{I},j\neq i} |a_{ij}||x_i| + |a_{ik}||x_i| + \sum_{j\in\mathcal{N}\backslash\mathcal{I},j\neq k} |a_{ij}||x_i| \\ & = & \sum_{j,j\neq i} |a_{ij}||x_i|, \end{array}$$

$|\lambda - a_{ii}| < \sum_{j,j\neq i} |a_{ij}|$, i.e. $\lambda \in C_i$.

Assume now $\mathcal{I} = \mathcal{N}$, that is all entries of the eigenvector $\mathbf{x}$ have the same absolute value. In this case:

$$|\lambda - a_{ii}||x_i| \leq \sum_{j,j\neq i} |a_{ij}||x_j| = \sum_{j,j\neq i} |a_{ij}||x_i|,\ \forall\, i,$$

$|\lambda - a_{ii}| \leq \sum_{j,j\neq i} |a_{ij}|$, $\forall\, i$, therefore either $\lambda \in C_s$ for some s or $\lambda \in \partial C_i\ \forall\, i$.

$\square$ Use the result obtained to prove that any irreducible weakly diagonal dominant $n \times n$ matrix A is non singular

$\square$ $\rho(A) \leq \|A\|_\infty$.

By the Gershgorin first theorem, for any eigenvalue λ of A there exists i such that $|\lambda| = |\lambda - a_{ii} + a_{ii}| \leq |\lambda - a_{ii}| + |a_{ii}| \leq \sum_j |a_{ij}| \leq \|A\|_\infty$

□ If A is irreducible and $\sum_j |a_{sj}| < \|A\|_\infty$ for some s, then $\rho(A) < \|A\|_\infty$.

Given an eigenvalue λ of A, the Gershgorin first theorem for irreducible matrices implies either $\exists\, i \mid |\lambda| = |\lambda - a_{ii} + a_{ii}| \leq |\lambda - a_{ii}| + |a_{ii}| < \sum_j |a_{ij}| \leq \|A\|_\infty$ or $|\lambda| = |\lambda - a_{ii} + a_{ii}| \leq |\lambda - a_{ii}| + |a_{ii}| = \sum_j |a_{ij}|$, $\forall\, i$, also for $i = s$, for which we know that $\sum_j |a_{sj}| < \|A\|_\infty$

(Jacobi method is able to solve linear systems $A\mathbf{x} = \mathbf{b}$ with A weakly diagonal dominant because in this case the Jacobi iteration matrix J satisfies the conditions $\exists\, s \mid \sum_j |[J]_{sj}| < \|J\|_\infty$ and $\|J\|_\infty = 1$, thus, by the result of the Exercise, $\rho(J) < 1$).

Proof of the existence of the SVD of $A \in \mathbb{C}^{n\times n}$

A $n \times n \Rightarrow \exists\, U, \sigma, V$, U, V unitary, $\sigma = \operatorname{diag}(\sigma_i)$ with $\sigma_1 \geq \sigma_2 \ldots \geq \sigma_n$ such that $A = U\sigma V^*$.

Proof. Let $\mathbf{v}_1$, $\|\mathbf{v}_1\|_2 = 1$, be such that $\|A\|_2 = \|A\mathbf{v}_1\|_2$ and set $\mathbf{u}_1 = A\mathbf{v}_1/\|A\mathbf{v}_1\|_2$ ($\|\mathbf{u}_1\|_2 = 1$ and $A\mathbf{v}_1 = \|A\|_2\mathbf{u}_1$). Let $\tilde{\mathbf{u}}_i, \tilde{\mathbf{v}}_i \in \mathbb{C}^n$ be such that $U = [\mathbf{u}_1|\tilde{\mathbf{u}}_2|\cdots|\tilde{\mathbf{u}}_n]$ and $V = [\mathbf{v}_1|\tilde{\mathbf{v}}_2|\cdots|\tilde{\mathbf{v}}_n]$ are unitary. Then

$$U^*AV = \begin{bmatrix} \mathbf{u}_1^* \\ \tilde{\mathbf{u}}_2^* \\ \cdots \\ \tilde{\mathbf{u}}_n^* \end{bmatrix} A[\mathbf{v}_1|\tilde{\mathbf{v}}_2|\cdots|\tilde{\mathbf{v}}_n] = \begin{bmatrix} \mathbf{u}_1^* \\ \tilde{\mathbf{u}}_2^* \\ \cdots \\ \tilde{\mathbf{u}}_n^* \end{bmatrix} [\|A\|_2\mathbf{u}_1|A\tilde{\mathbf{v}}_2|\cdots|A\tilde{\mathbf{v}}_n] = \begin{bmatrix} \|A\|_2 & \mathbf{w}^* \\ \mathbf{0} & \hat{A} \end{bmatrix},$$

$$\begin{aligned} \|A\|_2 &= \|U^*AV\|_2 = \sup_{\mathbf{v}\neq\mathbf{0}} \frac{\| \begin{bmatrix} \|A\|_2 & \mathbf{w}^* \\ \mathbf{0} & \hat{A} \end{bmatrix} \mathbf{v}\|_2}{\|\mathbf{v}\|_2} \\ &\geq \frac{\| \begin{bmatrix} \|A\|_2 & \mathbf{w}^* \\ \mathbf{0} & \hat{A} \end{bmatrix} \begin{bmatrix} \|A\|_2 \\ \mathbf{w} \end{bmatrix} \|_2}{\| \begin{bmatrix} \|A\|_2 \\ \mathbf{w} \end{bmatrix} \|_2} \geq \frac{\|A\|_2^2 + \|\mathbf{w}\|_2^2}{\sqrt{\|A\|_2^2 + \|\mathbf{w}\|_2^2}} = \sqrt{\|A\|_2^2 + \|\mathbf{w}\|_2^2} \end{aligned}$$

$\Rightarrow\ \ \mathbf{w} = \mathbf{0}\ \ \Rightarrow$

$$\begin{aligned} \|A\|_2 &= \|U^*AV\|_2 = \sup_{\mathbf{v}\neq\mathbf{0}} \frac{\| \begin{bmatrix} \|A\|_2 & \mathbf{0}^* \\ \mathbf{0} & \hat{A} \end{bmatrix} \mathbf{v}\|_2}{\|\mathbf{v}\|_2} \\ &\geq \sup_{\hat{\mathbf{v}}\neq\mathbf{0}} \frac{\| \begin{bmatrix} \|A\|_2 & \mathbf{0}^* \\ \mathbf{0} & \hat{A} \end{bmatrix} \begin{bmatrix} 0 \\ \hat{\mathbf{v}} \end{bmatrix} \|_2}{\| \begin{bmatrix} 0 \\ \hat{\mathbf{v}} \end{bmatrix} \|_2} = \|\hat{A}\|_2 \end{aligned}$$

$\Rightarrow\ U^*AV = \begin{bmatrix} \|A\|_2 & \mathbf{0}^* \\ \mathbf{0} & \hat{A} \end{bmatrix}$ with $\hat{A}$ such that $\|\hat{A}\|_2 \leq \|A\|_2$.

The thesis follows if we assume it true for matrices of order $n-1$.

On SVD: best rank-r approximation of A.

A $n \times n$, $A = U\sigma V^* = \sum_1^n \sigma_i \mathbf{u}_i \mathbf{v}_i^*$, $A_r = \sum_1^r \sigma_i \mathbf{u}_i \mathbf{v}_i^* \Rightarrow$

$$\min\{\|A - B\|_2 :\ \operatorname{rank}(B) \leq r\} = \|A - A_r\|_2 = \sigma_{r+1}$$

Proof. Let B be a $n \times n$ matrix with complex entries whose rank is no more than r and set $\mathcal{L} = \{\mathbf{v} :\ B\mathbf{v} = \mathbf{0}\}$. Observe that

$$\|A - B\|_2 = \sup_{\mathbf{v}} \frac{\|(A-B)\mathbf{v}\|_2}{\|\mathbf{v}\|_2} \geq \sup_{\mathbf{v} \in \mathcal{L}} \frac{\|A\mathbf{v}\|_2}{\|\mathbf{v}\|_2}.$$

Set $\mathcal{M} =$ Span $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_{r+1}\}$. Since $\dim \mathcal{M} + \dim \mathcal{L} \geq n+1$, there exists $\mathbf{z} \neq \mathbf{0}$, $\mathbf{z} \in \mathcal{M} \cap \mathcal{L}$,

$$\|A - B\|_2 \geq \frac{\|A\mathbf{z}\|_2}{\|\mathbf{z}\|_2} \geq \sigma_{r+1}$$

(first: $\mathbf{z} \in \mathcal{L}$; second: $\mathbf{z} \in \mathcal{M} \Rightarrow \mathbf{z} = \sum_1^{r+1} \alpha_i \mathbf{v}_i \Rightarrow A\mathbf{z} = \sum_1^{r+1} \alpha_i \sigma_i \mathbf{u}_i$). Moreover,

$$\|A - A_r\|_2 = \|U \operatorname{diag}(0, \ldots, 0, \sigma_{r+1}, \ldots, \sigma_n) V^*\|_2 = \|\operatorname{diag}(\ldots)\|_2 = \sigma_{r+1}$$

and $\operatorname{rank}(A_r) \leq r$.

Remark. We also have:

$$\min\{\|A - B\|_F :\ \operatorname{rank}(B) \leq r\} = \|A - A_r\|_F = \sqrt{\sum_{r+1}^{n} \sigma_j^2}$$

In functional analysis for compact operators ... (linear banded operators on Hilbert spaces) use * as a definition of singular values, approximate an object with something of finite dimension

---

On SVD: kernel and image of A.

A $n \times n$, $A = U\sigma V^*$, $\sigma_1 \geq \ldots \geq \sigma_k > 0 = \sigma_{k+1} = \ldots = \sigma_n \Rightarrow$

(1) $\{\mathbf{x} \in \mathbb{C}^n :\ A\mathbf{x} = \mathbf{0}\} =$ Span $\{\mathbf{v}_{k+1}, \ldots, \mathbf{v}_n\}$

(2) $\{A\mathbf{x} : \mathbf{x} \in \mathbb{C}^n\} =$ Span $\{\mathbf{u}_1, \ldots, \mathbf{u}_k\}$

(3) $\operatorname{rank}(A) = k = \#\{\sigma_i :\ \sigma_i > 0\}$

Proof. (1): $A\mathbf{x} = \mathbf{0}$ iff $\sigma V^*\mathbf{x} = \mathbf{0}$ iff $S_k V_k^* \mathbf{x} = \mathbf{0}$,

$$S_k = \begin{bmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_k \end{bmatrix}, \quad V_k = \begin{bmatrix} \mathbf{v}_1^* \\ \cdots \\ \mathbf{v}_k^* \end{bmatrix},$$

iff $V_k^*\mathbf{x} = \mathbf{0}$ iff $\mathbf{x}$ is orthogonal to $\mathbf{v}_1, \ldots, \mathbf{v}_k$ iff $\mathbf{x}$ is a linear combination of $\mathbf{v}_{k+1}, \ldots, \mathbf{v}_n$.

(2):

$$A\mathbf{x} = [U_k\ \Box] \begin{bmatrix} S_k & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} V_k^* \\ \Box \end{bmatrix} = U_k(S_k V_k^* \mathbf{x}),\ U_k = [\mathbf{u}_1 \cdots \mathbf{u}_k]$$

$\Rightarrow A\mathbf{x} \in$ Span $\{\mathbf{u}_1, \ldots, \mathbf{u}_k\} \Rightarrow \{A\mathbf{x} : \mathbf{x} \in \mathbb{C}^n\} \subset$ Span $\{\mathbf{u}_1, \ldots, \mathbf{u}_k\}$. Now let us show that for any $\mathbf{z} \in \mathbb{C}^k$ there exists $\mathbf{x} \in \mathbb{C}^n$, $U_k\mathbf{z} = A\mathbf{x}$:

$$\begin{aligned} &\exists\, \mathbf{x} \,|\, A\mathbf{x} = U_k\mathbf{z} \ \text{ iff} \\ &\exists\, \mathbf{x} \,|\, U_k S_k V_k^* \mathbf{x} = U_k \mathbf{z} \ \text{ iff} \\ &\exists\, \mathbf{x} \,|\, S_k V_k^* \mathbf{x} = \mathbf{z} \ \text{ iff} \\ &\exists\, \mathbf{x} \,|\, V_k^* \mathbf{x} = S_k^{-1} \mathbf{z}. \end{aligned}$$

Since $\text{rank}(V_k^*) = k$, the latter system admits solution.

On SVD: exercises

☐

$$A = \frac{1}{81}\begin{bmatrix} -65 & 76 & 104 \\ 76 & -206 & 8 \\ 104 & 8 & 109 \end{bmatrix} = UDU^*,$$

$$U = \frac{1}{9}\begin{bmatrix} -4 & 4 & 7 \\ 8 & 1 & 4 \\ 1 & 8 & -4 \end{bmatrix}, \quad D = \begin{bmatrix} -3 & & \\ & 2 & \\ & & -1 \end{bmatrix}.$$

Write the SVD of A.

☐ λ_i eigenvalues of $A \Rightarrow \sigma_n \le |\lambda_i| \le \sigma_1$.

($A\mathbf{x} = \lambda\mathbf{x}$, $A = U\sigma V^* \Rightarrow \mathbf{y}^*\sigma^2\mathbf{y} = \mathbf{x}^*A^*A\mathbf{x} = |\lambda|^2\|\mathbf{x}\|_2^2 \ldots$).

On SVD: how to compute the rank of a matrix, Gram-Schmidt vs SVD

Let $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_m, \ldots$ be a sequence of non null $n \times 1$ vectors and set $A_m = [\mathbf{a}_1\, \mathbf{a}_2 \cdots \mathbf{a}_m]$, $m = 1, 2, \ldots$. There follows an algorithm which computes matrices $Q_m = [\mathbf{q}_1\, \mathbf{q}_2 \cdots \mathbf{q}_m]$, $n \times m$, and R_m, upper triangular $m \times m$, such that

(1) $A_m = Q_m R_m$, $m = 1, 2, \ldots$

(2) $\{\mathbf{q}_1\} \cup \{\mathbf{q}_k : \ 2 \le k \le m, \mathbf{a}_k \notin \text{Span}\{\mathbf{a}_1, \ldots, \mathbf{a}_{k-1}\}\}$ is an orthonormal basis of the space $\text{Span}\{\mathbf{a}_1, \ldots, \mathbf{a}_m\}$

(3) if $\mathbf{a}_k$, $2 \le k \le m$ is linearly dependent from $\mathbf{a}_1, \ldots, \mathbf{a}_{k-1}$, then the k-row of R_m is null and $\mathbf{q}_k$ can be chosen arbitrarily (for instance, $\mathbf{q}_k = \mathbf{0}$ or such that $Q_m^* Q_m = I$)

(4) The rank of A_m is the number of non null rows of R_m

Set $\hat{\mathbf{q}}_1 = \mathbf{a}_1$ and $\mathbf{q}_1 = \hat{\mathbf{q}}_1/\|\hat{\mathbf{q}}_1\|_2$. Then $\mathbf{a}_1 = \|\hat{\mathbf{q}}_1\|_2\mathbf{q}_1$, i.e.

$$\begin{bmatrix}\mathbf{a}_1\end{bmatrix} = \begin{bmatrix}\mathbf{q}_1\end{bmatrix}[\|\hat{\mathbf{q}}_1\|_2].$$

Set $\hat{\mathbf{q}}_2 = \mathbf{a}_2 - r_{12}\mathbf{q}_1$, r_{12} such that $\mathbf{q}_1^*\hat{\mathbf{q}}_2 = 0$ ($r_{12} = \mathbf{q}_1^*\mathbf{a}_2$) and, if $\hat{\mathbf{q}}_2 \ne \mathbf{0}$, $\mathbf{q}_2 = \hat{\mathbf{q}}_2/\|\hat{\mathbf{q}}_2\|_2$. Then $\mathbf{a}_2 = r_{12}\mathbf{q}_1 + \|\hat{\mathbf{q}}_2\|_2\mathbf{q}_2$, i.e.

$$\begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 \end{bmatrix} = \begin{bmatrix} \mathbf{q}_1 & \mathbf{q}_2 \end{bmatrix}\begin{bmatrix} \|\hat{\mathbf{q}}_1\|_2 & r_{12} \\ 0 & \|\hat{\mathbf{q}}_2\|_2 \end{bmatrix}.$$

Else, if $\hat{\mathbf{q}}_2 = \mathbf{0}$, or, equivalently, $\mathbf{a}_2 = r_{12}\mathbf{q}_1 \in \text{Span}\{\mathbf{a}_1\}$, we can write

$$\begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 \end{bmatrix} = \begin{bmatrix} \mathbf{q}_1 & \mathbf{q}_2 \end{bmatrix}\begin{bmatrix} \|\hat{\mathbf{q}}_1\|_2 & r_{12} \\ 0 & 0 \end{bmatrix}, \ \mathbf{q}_2 := \hat{\mathbf{q}}_2 = \mathbf{0} \text{ or arbitrary.}$$

Assume that the first case occurs. Set $\hat{\mathbf{q}}_3 = \mathbf{a}_3 - r_{13}\mathbf{q}_1 - r_{23}\mathbf{q}_2$, r_{13}, r_{23} such that $\mathbf{q}_1^*\hat{\mathbf{q}}_3 = \mathbf{q}_2^*\hat{\mathbf{q}}_3 = 0$ ($r_{13} = \mathbf{q}_1^*\mathbf{a}_3$, $r_{23} = \mathbf{q}_2^*\mathbf{a}_3$) and assume $\hat{\mathbf{q}}_3 = \mathbf{0}$, or, equivalently,

$\mathbf{a}_3 = r_{13}\mathbf{q}_1 + r_{23}\mathbf{q}_2 \in \text{Span}\{\mathbf{a}_1, \mathbf{a}_2\}$. Then we can write:

$$\begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \mathbf{a}_3 \end{bmatrix} = \begin{bmatrix} \mathbf{q}_1 & \mathbf{q}_2 & \mathbf{q}_3 \end{bmatrix} \begin{bmatrix} \|\hat{\mathbf{q}}_1\|_2 & r_{12} & r_{13} \\ 0 & \|\hat{\mathbf{q}}_2\|_2 & r_{23} \\ 0 & 0 & 0 \end{bmatrix},$$
$$\mathbf{q}_3 := \hat{\mathbf{q}}_3 = \mathbf{0} \text{ or arbitrary.}$$

Set $\hat{\mathbf{q}}_4 = \mathbf{a}_4 - r_{14}\mathbf{q}_1 - r_{24}\mathbf{q}_2$, r_{14}, r_{24} such that $\mathbf{q}_1^*\hat{\mathbf{q}}_4 = \mathbf{q}_2^*\hat{\mathbf{q}}_4 = 0$ ($r_{14} = \mathbf{q}_1^*\mathbf{a}_4$, $r_{24} = \mathbf{q}_2^*\mathbf{a}_4$) and assume $\hat{\mathbf{q}}_4 = \mathbf{0}$, or, equivalently, $\mathbf{a}_4 = r_{14}\mathbf{q}_1 + r_{24}\mathbf{q}_2 \in \text{Span}\{\mathbf{a}_1, \mathbf{a}_2\}$. Then we can write:

$$\begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \mathbf{a}_3 & \mathbf{a}_4 \end{bmatrix} = \begin{bmatrix} \mathbf{q}_1 & \mathbf{q}_2 & \mathbf{q}_3 & \mathbf{q}_4 \end{bmatrix} \begin{bmatrix} \|\hat{\mathbf{q}}_1\|_2 & r_{12} & r_{13} & r_{14} \\ 0 & \|\hat{\mathbf{q}}_2\|_2 & r_{23} & r_{24} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix},$$
$$\mathbf{q}_3 := \hat{\mathbf{q}}_3 = \mathbf{0},\ \mathbf{q}_4 := \hat{\mathbf{q}}_4 = \mathbf{0} \text{ or arbitrary.}$$

Set $\hat{\mathbf{q}}_5 = \mathbf{a}_5 - r_{15}\mathbf{q}_1 - r_{25}\mathbf{q}_2$, r_{15}, r_{25} such that $\mathbf{q}_1^*\hat{\mathbf{q}}_5 = \mathbf{q}_2^*\hat{\mathbf{q}}_5 = 0$ ($r_{15} = \mathbf{q}_1^*\mathbf{a}_5$, $r_{25} = \mathbf{q}_2^*\mathbf{a}_5$) and assume $\hat{\mathbf{q}}_5 \neq \mathbf{0}$. Set $\mathbf{q}_5 = \hat{\mathbf{q}}_5/\|\hat{\mathbf{q}}_5\|_2$. Then $\mathbf{a}_5 = r_{15}\mathbf{q}_1 + r_{25}\mathbf{q}_2 + \|\hat{\mathbf{q}}_5\|_2\mathbf{q}_5$, i.e.

$$\begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \mathbf{a}_3 & \mathbf{a}_4 & \mathbf{a}_5 \end{bmatrix} = \begin{bmatrix} \mathbf{q}_1 & \mathbf{q}_2 & \mathbf{q}_3 & \mathbf{q}_4 & \mathbf{q}_5 \end{bmatrix} \begin{bmatrix} \|\hat{\mathbf{q}}_1\|_2 & r_{12} & r_{13} & r_{14} & r_{15} \\ 0 & \|\hat{\mathbf{q}}_2\|_2 & r_{23} & r_{24} & r_{25} \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \|\hat{\mathbf{q}}_5\|_2 \end{bmatrix},$$
$$\mathbf{q}_3, \mathbf{q}_4 \text{ null or arbitrary.}$$

...

Remark. Since the calculator uses finite arithmetic, the check if $\hat{\mathbf{q}}_k$, $k \geq 2$, is zero or nonzero must be replaced with something of type: $\|\hat{\mathbf{q}}_k\|$ is less than ε or not? Moreover, take into account that even a very little perturbation in one entry of a triangular matrix can change the value of its rank (see the following example). These facts imply that the (Gram-Schmidt) algorithm illustrated above may generate a *numeric rank* of A_m which is different from the rank of A_m.

Example. Let R be the $n \times n$ upper triangular matrix

$$R = \begin{bmatrix} 1 & -1 & -1 & \cdots & -1 \\ 0 & 1 & -1 & \cdots & -1 \\ \vdots & \ddots & 1 & \ddots & \vdots \\ \vdots & & \ddots & \ddots & -1 \\ 0 & \cdots & \cdots & 0 & 1 \end{bmatrix}.$$

The rank of R is n, but if the 0 in the $(n, 1)$ entry is replaced with -2^{2-n} (which for large n is a very little perturbation), then the rank of R becomes $n-1$. The SVD of R predicts this observation. In fact, the singular value σ_{n-1} of R for $n = 5, 10, 15$ has more or less the same value, 1.5, whereas the smallest singular value, σ_n, seems to tend to zero:

$$n = 5: \ \sigma_5 \approx \frac{1}{10}, \quad n = 10: \ \sigma_{10} \approx \frac{1}{100}, \quad n = 15: \ \sigma_{15} \approx \frac{1}{10000}.$$

So, by examining the singular values of R we see that even if $\det(R) = 1$ (far from zero) for all n, greater is n, smaller is the distance of R from a singular matrix. (Note that R is not normal, in fact $\mu_2(R) = \sigma_1/\sigma_n \approx 30, 2000, 10^5 > 1 = \max|\lambda_i|/\min|\lambda_i|$).

It is known that small perturbations on the entries of A imply at most small perturbations on U, σ, V, $A = U\sigma V^*$ (SVD problem is well conditioned). It follows that the algorithm for the computation of the SVD of A can give accurate approximations of U, σ, V. Having an accurate approximation of σ we can evaluate precisely the rank of A; we can even quantify how much A is far from having a smaller rank. Thus it is preferable to compute the rank of a matrix via SVD, instead via Gram-Schmidt.