1. We extend the orthogonal transforms in QR decomposition algorithms to complex matrices.

(a) Find the complex Householder matrix. Given a nonzero $\boldsymbol{v} \in \mathbb{C}^n$, find a matrix $\boldsymbol{H} \in \mathbb{C}^{n\times n}$ such that $\boldsymbol{H}\boldsymbol{H}^* = \boldsymbol{H}^*\boldsymbol{H} = \boldsymbol{I}$ and $\boldsymbol{H}\boldsymbol{v}$ is a multiple of $\boldsymbol{e}_1$, where the superscript $*$ stands for conjugate transpose.

(b) Find the complex Givens rotation matrix. More precisely, given $\boldsymbol{x} \in \mathbb{C}^2$, find a matrix $\boldsymbol{G} \in \mathbb{C}^{2\times 2}$ such that $\boldsymbol{G}\boldsymbol{G}^* = \boldsymbol{G}^*\boldsymbol{G} = \boldsymbol{I}$ and $\boldsymbol{G}\boldsymbol{x} = \begin{bmatrix} \times \\ 0 \end{bmatrix}$, where the superscript $*$ stands for conjugate transpose and $\times$ is a complex number.

(a) For $v = (v_1, v_2, \cdots, v_n)^T \in \mathbb{C}^n$, let $H_u = -e^{-i\theta}(I - 2uu^*)$.

here u is unit of $v + e^{i\theta}\|v\|_2 e_1$, $e^{i\theta} = \frac{v_1}{|v_1|}$.

Then

① $H_u H_u^* = (I - 2uu^*)(I - 2uu^*)$

$$= I - 4uu^* + 4u(u^*u)u^* \quad (\text{by } u \text{ is unit})$$

$$= I - 4uu^* + 4uu^*$$

$$= I$$

and also $H_u^* H_u = I$

② $H_u v = -e^{-i\theta}(I - 2uu^*)v$

$$= -e^{-i\theta}v + 2e^{-i\theta}uu^*v$$

$$= -e^{-i\theta}v + 2\frac{e^{-i\theta}v + \|v\|_2 e_1}{\|v + e^{i\theta}\|v\|_2 e_1\|_2^2}(\|v\|_2^2 + |v_1|\,\|v\|_2)$$

$$\|v + e^{i\theta}\|v\|_2 e_1\|_2^2 = v^*v + e^{i\theta}\|v\|_2 v^* e_1 + e^{-i\theta}\|v\|_2 e_1^* v + \|v\|_2^2 e_1^* e_1$$

$$= 2\|v\|_2^2 + 2|v_1|\,\|v\|_2$$

hence $H_u v = -e^{-i\theta}v + e^{-i\theta}v + \|v\|_2 e_1 = \|v\|_2 e_1$

(b) Let $G = \begin{pmatrix} c & s \\ -\bar{s} & \bar{c} \end{pmatrix}$, then G satisfies $GG^* = G^*G$.

$$GG^* = \begin{pmatrix} c & s \\ -\bar{s} & \bar{c} \end{pmatrix}\begin{pmatrix} \bar{c} & -s \\ \bar{s} & c \end{pmatrix} = \begin{pmatrix} c\bar{c} + s\bar{s} & 0 \\ 0 & s\bar{s} + c\bar{c} \end{pmatrix} = I.$$

We have $|c|^2 + |s|^2 = 1$.

For $x = (x_1, x_2)^T \in \mathbb{C}^2$, $Gx = \begin{pmatrix} cx_1 + sx_2 \\ -\bar{s}x_1 + \bar{c}x_2 \end{pmatrix} = \begin{pmatrix} X \\ 0 \end{pmatrix}$,

then $-\bar{s}x_1 + \bar{c}x_2 = 0$. We choose

$$c = \frac{\bar{x}_1}{\|x\|_2}, \quad s = \frac{\bar{x}_2}{\|x\|_2}$$

and $G = \begin{pmatrix} \bar{x}_1/\|x\|_2 & \bar{x}_2/\|x\|_2 \\ -x_2/\|x\|_2 & x_1/\|x\|_2 \end{pmatrix}$.

2. Consider the least squares problem $\min_{\boldsymbol{x}\in\mathbb{R}^n} \|\boldsymbol{A}\boldsymbol{x}-\boldsymbol{b}\|_2^2$, where $\boldsymbol{A} \in \mathbb{R}^{m\times n}$ with $m > n$ is of full column rank.

(a) Prove: $\boldsymbol{x}$ is the solution of the least squares problem if and only if there exists a vector $\boldsymbol{y} \in \mathbb{R}^m$ such that

$$\begin{bmatrix} \boldsymbol{I} & \boldsymbol{A} \\ \boldsymbol{A}^T & \boldsymbol{0} \end{bmatrix} \begin{bmatrix} \boldsymbol{y} \\ \boldsymbol{x} \end{bmatrix} = \begin{bmatrix} \boldsymbol{b} \\ \boldsymbol{0} \end{bmatrix} \tag{1}$$

This provides another way to solve the least squares problem.

(b) Prove that the coefficient matrix in (1) is non-singular and symmetric indefinite. (A matrix is indefinite if the associated quadratic form can be either positive or negative.)

(c) Prove that, if $\boldsymbol{A}$ is properly scaled such that the minimum eigenvalue of $\boldsymbol{A}^T\boldsymbol{A}$ is 1, then the condition number of the coefficient matrix in (1) is upper bounded by $C \cdot \sqrt{\mathrm{Cond}(\boldsymbol{A}^T\boldsymbol{A})}$ for some universal constant $C > 0$, where $\mathrm{Cond}(\cdot)$ is the condition number. Therefore, (1) is better to solve than the normal equation, though it is indefinite.

(a) Since $A \in \mathbb{R}^{m\times n}$ and $m>n$, $\mathrm{Null}(A^T)$ is nontrivial.

x is the solution of $\min_{x\in\mathbb{R}^n} \|Ax-b\|_2^2$

$\Leftrightarrow b - Ax \perp \mathrm{Ran}(A)$

$\Leftrightarrow \exists y \in \mathbb{R}^m$ s.t. $b-Ax=y$ and $y \in \mathrm{Ran}(A)^\perp = \mathrm{Null}(A^T)$

$\Leftrightarrow \begin{pmatrix} I & A \\ A^T & 0 \end{pmatrix}\begin{pmatrix} y \\ x \end{pmatrix} = \begin{pmatrix} b \\ 0 \end{pmatrix}$

(b) ① Since $\begin{pmatrix} I & 0 \\ -A^T & I \end{pmatrix}\begin{pmatrix} I & A \\ A^T & 0 \end{pmatrix}\begin{pmatrix} I & -A \\ 0 & I \end{pmatrix} = \begin{pmatrix} I & 0 \\ 0 & -A^TA \end{pmatrix}$.

$\det\begin{pmatrix} I & A \\ A^T & 0 \end{pmatrix} = \det(-A^TA) \neq 0$, as A is of full column rank. So $\begin{pmatrix} I & A \\ A^T & 0 \end{pmatrix}$ is non-singular.

② $\begin{pmatrix} I & A \\ A^T & 0 \end{pmatrix}$ is symmetric.

③ $(y^T \ x^T)\begin{pmatrix} I & A \\ A^T & 0 \end{pmatrix}\begin{pmatrix} y \\ x \end{pmatrix} = y^Ty + 2y^TAx$.

If we pick $y \in \mathrm{Null}(A^T)$ and $y \neq 0$,
then $y^Ty + 2y^TAx = \|y\|^2 > 0$.

If we pick x such that $Ax \neq 0$ and $y = -Ax$,
then $y^Ty + 2y^TAx = -(Ax)^T(Ax) < 0$.

Hence, $\begin{pmatrix} I & A \\ A^T & 0 \end{pmatrix}$ is symmetric indefinite.

(c) Let $A = U\Sigma V^*$ the SVD of A.

U and V are unitary, $\Sigma = \begin{pmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_n \\ & 0 & \end{pmatrix}_{m\times n}$, $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_n = 1$

Since $\begin{pmatrix} U & \\ & V \end{pmatrix}\begin{pmatrix} I & \Sigma \\ \Sigma^T & 0 \end{pmatrix}\begin{pmatrix} U^* & \\ & V^* \end{pmatrix} = \begin{pmatrix} I & A \\ A^T & 0 \end{pmatrix}$,

the eigenvalues of $\begin{pmatrix} I & \Sigma \\ \Sigma^T & 0 \end{pmatrix}$ is the same as $\begin{pmatrix} I & A \\ A^T & 0 \end{pmatrix}$.

Now we consider the eigenvalues of $\begin{pmatrix} I & \Sigma \\ \Sigma^T & 0 \end{pmatrix}$.

$$\det\begin{pmatrix} \lambda I - I & -\Sigma \\ -\Sigma^T & \lambda I \end{pmatrix} = \frac{1}{\lambda^m}\det\left[\begin{pmatrix} (\lambda-1)I & -\Sigma \\ -\Sigma^T & \lambda I \end{pmatrix}\begin{pmatrix} \lambda I & 0 \\ \Sigma^T & I \end{pmatrix}\right]$$

$$= \frac{1}{\lambda^m}\det\begin{pmatrix} (\lambda-1)\lambda I - \Sigma\Sigma^T & X \\ 0 & \lambda I \end{pmatrix}$$

$$= \lambda^{n-m}\det\left((\lambda-1)\lambda I - \Sigma\Sigma^T\right)$$

$$= \lambda^{n-m}\det\begin{pmatrix} (\lambda-1)\lambda - \sigma_1^2 & & & & \\ & \ddots & & & \\ & & (\lambda-1)\lambda - \sigma_n^2 & & \\ & & & (\lambda-1)\lambda & \\ & & & & (\lambda-1)\lambda \end{pmatrix}$$

$$= \lambda^{n-m}\left[(\lambda-1)\lambda\right]^{m-n}\prod_{k=1}^{n}\left[(\lambda-1)\lambda - \sigma_k^2\right]$$

$$= (\lambda-1)^{m-n}\prod_{k=1}^{n}\left[(\lambda-1)\lambda - \sigma_k^2\right]$$

Thus the eigenvalues of $\begin{pmatrix} I & \Sigma \\ \Sigma^T & 0 \end{pmatrix}$ are $1, \frac{1}{2}\left(1 \pm \sqrt{1+4\sigma_k^2}\right)$

$k = 1, 2, \cdots, n$

$$|\lambda|_{max} = \frac{1}{2}\left(1 + \sqrt{1+4\sigma_1^2}\right)$$

$$|\lambda|_{min} = \frac{1}{2}\left(\sqrt{1+4\sigma_n^2} - 1\right) = \frac{\sqrt{5}-1}{2}$$

$$\mathrm{Cond}\begin{pmatrix} I & A \\ A^T & 0 \end{pmatrix} = \mathrm{Cond}\begin{pmatrix} I & \Sigma \\ \Sigma^T & 0 \end{pmatrix}$$ by similarity and symmetry

$$= \frac{1+\sqrt{1+4\sigma_1^2}}{\sqrt{5}-1}$$

$$\leq \frac{\sqrt{5}+1}{\sqrt{5}-1}\sigma_1$$ by $\sigma_1 \geq \sigma_n = 1$

$$= \frac{\sqrt{5}+1}{\sqrt{5}-1}\sqrt{\mathrm{Cond}(A^TA)}$$

3. The QR algorithm can also be used in theoretical analysis. Let $\boldsymbol{A} = [\boldsymbol{a}_1\ \boldsymbol{a}_2\ \dots\ \boldsymbol{a}_n] \in \mathbb{R}^{n\times n}$ with column vectors $\boldsymbol{a}_i$ for $i = 1, \dots, n$. Prove

$$|\det(\boldsymbol{A})| \le \prod_{i=1}^{n} \|\boldsymbol{a}_i\|_2,$$

and the upper bound is attained if and only if $\boldsymbol{A}^T\boldsymbol{A}$ is diagonal or $\|\boldsymbol{a}_i\|_2 = 0$ for some i. (*Hint: Consider the Gram-Schmidt procedure.*)

Consider QR via Gram-Schmidt procedure. $A = QR$

where $Q = (q_1, q_2, \cdots q_n)$ is orthogonal and

R is upper triangular.

$$
\begin{aligned}
|\det A| &= |\det Q|\,|\det R| \\
&= \prod_{i=1}^{n} |r_{ii}| \\
&= \prod_{i=1}^{n} |q_i^T a_i| \\
&\le \prod_{i=1}^{n} \|q_i\|_2 \|a_i\|_2 \\
&= \prod_{i=1}^{n} \|a_i\|_2
\end{aligned}
$$

the upper bound is attained

$$\Leftrightarrow \prod_{i=1}^{n} |q_i^T a_i| = \prod_{i=1}^{n} \|q_i\|_2 \|a_i\|_2 \qquad (*)$$

$\Leftrightarrow$ ① If for some i such that $\|a_i\|_2 = 0$

or

② If for all $i \in \{1, \cdots, n\}$, $\|a_i\|_2 \ne 0$, then

$$|q_i^T a_i| = \|q_i\|_2 \|a_i\|_2 \qquad i = 1, 2, \cdots, n$$

$$\Leftrightarrow \quad a_i = d_i q_i \qquad i = 1, 2, \cdots, n$$

$$\Leftrightarrow \quad A = QD, \quad D = \begin{pmatrix} d_1 & & \\ & \ddots & \\ & & d_n \end{pmatrix}$$

$$\Leftrightarrow \quad A^T A \text{ is diagonal}$$

4. Let the initial vector $\boldsymbol{x}_0 = \boldsymbol{0}$. When the GMRES is applied to solve $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$, where

$$\boldsymbol{A} = \begin{bmatrix} 0 & & & & 1 \\ 1 & 0 & & & \\ & 1 & 0 & & \\ & & \ddots & \ddots & \\ & & & 1 & 0 \end{bmatrix}, \qquad \boldsymbol{b} = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix},$$

what is the convergence rate?

Given $x_0 = 0$. $r_0 = b - Ax_0 = b$

$$K_k = \text{span}\{r_0, Ar_0, \cdots, A^{k-1}r_0\}$$

$$= \text{span}\{e_1, Ae_1, \cdots, A^{k-1}e_1\}$$

$$= \text{span}\{e_1, e_2, \cdots, e_k\}$$

$$x_k = \underset{x \in x_0 + K_k}{\text{argmin}} \|Ax - b\|_2$$

$$\|r_k\|_2 = \min_{y \in \mathbb{R}^k} \|(b - Ax_0) - \sum_{j=1}^{k} y_j A^j r_0\|_2$$

$$= \min_{y \in \mathbb{R}^k} \left\| \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} - \begin{pmatrix} 0 \\ y_1 \\ \vdots \\ y_k \\ 0 \\ \vdots \\ 0 \end{pmatrix} \right\|_2$$

$$= 1$$

For $k = 1, 2, \cdots, n-1$, $\|r_k\|_2 = 1$ and when $k = n$, $\|r_n\|_2 = 0$.

then the convergence rate is $\dfrac{\|r_k\|_2}{\|r_{k-1}\|_2} = 1$, and it will converge

at $k = n$.