APPM 5610 - Homework 1

Eappen Nelluvelil

January 26, 2022

1. Any square matrix may be represented in the form $\mathbf{A} = \mathbf{S}\mathbf{V}$, where $\mathbf{S}$ is a Hermitian nonnegative definite matrix and $\mathbf{V}$ is a unitary matrix. This is the so-called polar decomposition of matrices (by analogy with the polar decomposition of complex numbers).

   Using the polar decomposition, derive the singular value decomposition for square matrices. Do not forgot to formulate explicitly what is the singular value decomposition of an arbitrary matrix.

   Recall that any matrix $\mathbf{A} \in \mathbb{C}^{m \times n}$ can be written as $\mathbf{A} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^*$, where

   (a) r is the rank of $\mathbf{A}$;
   (b) $\mathbf{U} \in \mathbb{C}^{m \times r}$ has orthonormal columns;
   (c) $\mathbf{\Sigma} \in \mathbb{R}^{r \times r} = \text{diag}\,(s_1, \ldots, s_r)$, where s_i is the i^{th} singular value of $\mathbf{A}$ for $1 \leq i \leq r$; and
   (d) $\mathbf{V} \in \mathbb{C}^{n \times r}$ has orthonormal columns.

   Let $\mathbf{A} \in \mathbb{C}^{n \times n}$ be given. We know that $\mathbf{A} = \mathbf{S}\widetilde{\mathbf{V}}$, where

   (a) $\mathbf{S}$ is Hermitian, semi-positive definite with rank r, and
   (b) $\widetilde{\mathbf{V}}$ is unitary.

   Since $\mathbf{S}$ is Hermitian, we know that $\mathbf{S}$ is diagonalizble. That is, $\mathbf{S} = \widetilde{\mathbf{U}}\widetilde{\mathbf{\Sigma}}\widetilde{\mathbf{U}}^*$, where

   (a) $\widetilde{\mathbf{U}}$ is unitary and whose columns are the eigenvectors of $\mathbf{S}$, and
   (b) $\widetilde{\mathbf{\Sigma}} = \text{diag}\,(\lambda_1, \ldots, \lambda_n)$, where $\lambda_i \geq 0$ is the i^{th} eigenvalue of $\mathbf{S}$.

   Define $\mathbf{U} = \begin{bmatrix} \widetilde{\mathbf{U}}_1 & \ldots & \widetilde{\mathbf{U}}_r \end{bmatrix}$, where $\mathbf{U} \in \mathbb{C}^{n \times r}$ and $\mathbf{U}^*\mathbf{U} = \mathbf{I}_{r \times r}$. We also define $\mathbf{\Sigma} = \text{diag}\,(\lambda_1, \ldots, \lambda_r)$, where $\mathbf{\Sigma} \in \mathbb{R}^{r \times r}$. Since the rank of $\mathbf{S}$ is r, by construction, $\mathbf{S} = \mathbf{U}\mathbf{\Sigma}\mathbf{U}^*$. Thus, we have that $\mathbf{A} = \mathbf{U}\mathbf{\Sigma}\mathbf{U}^*\widetilde{\mathbf{V}}$. We now define $\mathbf{V} = \widetilde{\mathbf{V}}^*\mathbf{U}$. By construction,

$$\begin{aligned} \mathbf{V}^*\mathbf{V} &= \mathbf{U}^*\widetilde{\mathbf{V}}\widetilde{\mathbf{V}}^*\mathbf{U} \\ &= \mathbf{U}^*\mathbf{U} \\ &= \mathbf{I}_{r \times r}, \end{aligned}$$

   i.e., $\mathbf{V}$ has orthonormal columns. Thus, $\mathbf{A} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^*$, which is the singular-value decomposition for square matrices.

2. For each of the following statements, prove that it is true or give an example to show that it is false. In all questions, $\mathbf{A} \in \mathbb{C}^{n \times n}$.

   (a) If $\mathbf{A}$ is real and λ is an eigenvalue of $\mathbf{A}$, then so is $-\lambda$.

   Consider the real matrix $\mathbf{A} = \begin{bmatrix} 1 & 0 \\ 0 & 4 \end{bmatrix}$. The eigenvalues of $\mathbf{A}$ are $\lambda_1 = 1$ and $\lambda_2 = 4$. By inspection, -1 and -4 are not eigenvalues of $\mathbf{A}$. Thus, this statement is false.

(b) If $\mathbf{A}$ is real and λ is an eigenvalue of $\mathbf{A}$, then so is $\overline{\lambda}$.

Suppose that $\mathbf{A}$ is real and $(\lambda, \mathbf{v})$ is an eigenpair of $\mathbf{A}$, i.e.,

$$\mathbf{A}\mathbf{v} = \lambda\mathbf{v}.$$

Taking the complex conjugate of both sides, we see that

$$\overline{\mathbf{A}\mathbf{v}} = \overline{\lambda\mathbf{v}} \implies \mathbf{A}\overline{\mathbf{v}} = \overline{\lambda}\overline{\mathbf{v}},$$

since $\mathbf{A} = \overline{\mathbf{A}}$. From this, we see that $\left(\overline{\lambda}, \overline{\mathbf{v}}\right)$ is also an eigenpair of $\mathbf{A}$. Thus, the statement is true.

(c) If λ is an eigenvalue of $\mathbf{A}$ and $\mathbf{A}$ is non-singular, then λ^{-1} is an eigenvalue of $\mathbf{A}^{-1}$.

Suppose that $(\lambda, \mathbf{v})$ is an eigenpair of $\mathbf{A}$, i.e., $\mathbf{A}\mathbf{v} = \lambda\mathbf{v}$, where $\mathbf{v} \in \mathbb{C}^n, \mathbf{v} \neq 0$, and also suppose that $\mathbf{A}^{-1}$ exists. Then, we have that

$$\begin{aligned} \mathbf{A}\mathbf{v} = \lambda\mathbf{v} &\implies \mathbf{A}^{-1}\mathbf{A}\mathbf{v} = \lambda\mathbf{A}^{-1}\mathbf{v} \\ &\implies \mathbf{A}^{-1}\mathbf{v} = \frac{1}{\lambda}\mathbf{v}. \end{aligned}$$

We see that $\frac{1}{\lambda}$ is an eigenvalue of $\mathbf{A}^{-1}$. Thus, the statement is true.

(d) If $\mathbf{A}$ is Hermitian and λ is an eigenvalue of $\mathbf{A}$, then $|\lambda|$ is a singular value of $\mathbf{A}$.

Let $\mathbf{A}$ be a given Hermitian matrix, i.e., $\mathbf{A} = \mathbf{A}^*$. Then, it follows that $\mathbf{A}$ is a normal matrix, i.e., $\mathbf{A}^*\mathbf{A} = \mathbf{A}\mathbf{A}^*$. Since $\mathbf{A}$ is normal, $\mathbf{A}$ is unitarily similar to a diagonal matrix. That is, $\mathbf{A} = \mathbf{U}\mathbf{\Sigma}\mathbf{U}^*$, where $\mathbf{U}$ is unitary and $\mathbf{\Sigma} = \text{diag}\left(\lambda_1, \ldots, \lambda_n\right) \in \mathbb{R}^{n\times n}$, with λ_i being the i^{th} eigenvalue of $\mathbf{A}$, $1 \leq i \leq n$. Note that if $(\lambda_i, \mathbf{v}_i)$ is an eigenpair of $\mathbf{A}$, we have that

$$\begin{aligned} \mathbf{A}\mathbf{v}_i = \lambda_i\mathbf{v}_i &\implies \mathbf{A}^*\mathbf{A}\mathbf{v}_i = \lambda_i\mathbf{A}^*\mathbf{v}_i \\ &\implies \mathbf{A}^2\mathbf{v}_i = \lambda_i^2\mathbf{v}_i, \end{aligned}$$

since $\mathbf{A}$ is Hermitian. The singular values of $\mathbf{A}$ are the square roots of the eigenvalues of $\mathbf{A}^*\mathbf{A}$, which we see are given by λ_i^2, $1 \leq i \leq n$. We see that σ_i, the i^{th} singular value of $\mathbf{A}$, is given by $\sigma_i = \sqrt{\lambda_i^2} = |\lambda_i|$. Thus, the statement is true.

3. A matrix $\mathbf{S} \in \mathbb{C}^{n\times n}$ such that $\mathbf{S}^* = -\mathbf{S}$ is called skew-Hermitian (recall that $\mathbf{S}^*$ denotes the adjoint matrix). Show that

(a) eigenvalues of $\mathbf{S}$ are pure imaginary (or zero).

Suppose that $(\lambda, \mathbf{v})$ is an eigenpair of $\mathbf{S}$, i.e., $\mathbf{S}\mathbf{v} = \lambda\mathbf{v}$. We see that

$$\begin{aligned} \langle \mathbf{S}\mathbf{v}, \mathbf{S}\mathbf{v} \rangle = \langle \lambda\mathbf{v}, \lambda\mathbf{v} \rangle &\implies \langle \mathbf{v}, \mathbf{S}^*\mathbf{S}\mathbf{v} \rangle = \langle \lambda\mathbf{v}, \lambda\mathbf{v} \rangle \\ &\implies \langle \mathbf{v}, -\mathbf{S}\mathbf{S}\mathbf{v} \rangle = |\lambda|^2 \, ||\mathbf{v}||_2^2 \\ &\implies -\,||\mathbf{S}\mathbf{v}||_2^2 = |\lambda|^2 \, ||\mathbf{v}||_2^2 \,. \end{aligned}$$

Since $||\mathbf{v}||_2^2 > 0$ and $||\mathbf{S}\mathbf{v}||_2^2 \geq 0$, this implies that λ is either purely imaginary or zero, as desired.

(b) matrix $\mathbf{I} - \mathbf{S}$ is non-singular.

For the sake of contradiction, suppose that $(\mathbf{I} - \mathbf{S})$ is singular, i.e., there exists a non-zero vector $\mathbf{x} \in \mathbb{C}^n$ such that $(\mathbf{I} - \mathbf{S})\,\mathbf{x} = \mathbf{0}$. From this, we see that

$$(\mathbf{I} - \mathbf{S})\,\mathbf{x} = \mathbf{0} \implies \mathbf{S}\mathbf{x} = \mathbf{x},$$

which implies that $\mathbf{x}$ is an eigenvector of $\mathbf{S}$ with corresponding eigenvalue $\lambda = 1$. However, this is a contradiction since we showed earlier that the eigenvalues of $\mathbf{S}$ are either purely imaginary or zero. Thus, it must be the case that $\mathbf{x} = \mathbf{0}$, which implies that $(\mathbf{I} - \mathbf{S})$ is non-singular, as desired.

(c) matrix $\mathbf{Q} = (\mathbf{I} - \mathbf{S})^{-1}(\mathbf{I} + \mathbf{S})$ is unitary (it is known as the Cayley transform of matrix $\mathbf{S}$).
Note that we can write $(\mathbf{I} + \mathbf{S}) = -(\mathbf{I} - \mathbf{S}) + 2\mathbf{I}$. We see that

$$\begin{aligned}(\mathbf{I}+\mathbf{S})(\mathbf{I}-\mathbf{S})^{-1} &= (-(\mathbf{I}-\mathbf{S})+2\mathbf{I})(\mathbf{I}-\mathbf{S})^{-1} \\ &= -(\mathbf{I}-\mathbf{S})(\mathbf{I}-\mathbf{S})^{-1} + 2(\mathbf{I}-\mathbf{S})^{-1} \\ &= -\mathbf{I} + 2(\mathbf{I}-\mathbf{S})^{-1},\end{aligned}$$

and we also see that

$$\begin{aligned}(\mathbf{I}-\mathbf{S})^{-1}(\mathbf{I}+\mathbf{S}) &= (\mathbf{I}-\mathbf{S})^{-1}(-(\mathbf{I}-\mathbf{S})+2\mathbf{I}) \\ &= -(\mathbf{I}-\mathbf{S})^{-1}(\mathbf{I}-\mathbf{S}) + 2(\mathbf{I}-\mathbf{S})^{-1} \\ &= -\mathbf{I} + 2(\mathbf{I}-\mathbf{S})^{-1}.\end{aligned}$$

Thus, the matrices $(\mathbf{I}+\mathbf{S})$ and $(\mathbf{I}-\mathbf{S})^{-1}$ commute. Taking $\mathbf{Q}$ as above, we see that

$$\begin{aligned}\mathbf{Q}^*\mathbf{Q} &= (\mathbf{I}+\mathbf{S})^*\left((\mathbf{I}-\mathbf{S})^{-1}\right)^*(\mathbf{I}-\mathbf{S})^{-1}(\mathbf{I}+\mathbf{S}) \\ &= (\mathbf{I}+\mathbf{S}^*)\left((\mathbf{I}-\mathbf{S})^*\right)^{-1}(\mathbf{I}-\mathbf{S})^{-1}(\mathbf{I}+\mathbf{S}) \\ &= (\mathbf{I}-\mathbf{S})\left((\mathbf{I}-\mathbf{S}^*)\right)^{-1}(\mathbf{I}-\mathbf{S})^{-1}(\mathbf{I}+\mathbf{S}) \\ &= (\mathbf{I}-\mathbf{S})(\mathbf{I}+\mathbf{S})^{-1}(\mathbf{I}-\mathbf{S})^{-1}(\mathbf{I}+\mathbf{S}) \\ &= (\mathbf{I}+\mathbf{S})^{-1}(\mathbf{I}-\mathbf{S})(\mathbf{I}-\mathbf{S})^{-1}(\mathbf{I}+\mathbf{S}) \\ &= (\mathbf{I}+\mathbf{S})^{-1}(\mathbf{I}+\mathbf{S}) \\ &= \mathbf{I}.\end{aligned}$$

We also see that

$$\begin{aligned}\mathbf{Q}\mathbf{Q}^* &= (\mathbf{I}-\mathbf{S})^{-1}(\mathbf{I}+\mathbf{S})(\mathbf{I}+\mathbf{S})^*\left((\mathbf{I}-\mathbf{S})^{-1}\right)^* \\ &= (\mathbf{I}-\mathbf{S})^{-1}(\mathbf{I}+\mathbf{S})(\mathbf{I}+\mathbf{S}^*)\left((\mathbf{I}-\mathbf{S})^*\right)^{-1} \\ &= (\mathbf{I}-\mathbf{S})^{-1}(\mathbf{I}+\mathbf{S})(\mathbf{I}-\mathbf{S})\left((\mathbf{I}-\mathbf{S}^*)\right)^{-1} \\ &= (\mathbf{I}-\mathbf{S})^{-1}(\mathbf{I}+\mathbf{S})(\mathbf{I}-\mathbf{S})(\mathbf{I}+\mathbf{S})^{-1} \\ &= (\mathbf{I}-\mathbf{S})^{-1}(\mathbf{I}+\mathbf{S})(\mathbf{I}+\mathbf{S})^{-1}(\mathbf{I}-\mathbf{S}) \\ &= (\mathbf{I}-\mathbf{S})^{-1}(\mathbf{I}-\mathbf{S}) \\ &= \mathbf{I}.\end{aligned}$$

Thus, $\mathbf{Q}$ is a unitary matrix, as desired.

4. Given $\mathbf{A} \in \mathbb{C}^{n\times n}$, use Schur's decomposition to show that, for every $\epsilon > 0$, there exists a diagonalizable matrix $\mathbf{B}$ such that $||\mathbf{A}-\mathbf{B}||_2 \le \epsilon$. This show that the set of diagonalizable matrices is dense in $\mathbb{C}^{n\times n}$ and that Jordan canonical form is not a continuous matrix decomposition.

   Let $\mathbf{A} \in \mathbb{C}^{n\times n}$ and $\epsilon > 0$ be given. By Schur's decomposition, $\mathbf{A}$ is unitarily similar to an upper triangular matrix, whose diagonal entries are the eigenvalues of $\mathbf{A}$. That is, $\mathbf{A} = \mathbf{U}\mathbf{T}\mathbf{U}^*$, where $\mathbf{T}$ is an upper triangular matrix whose diagonal entries are the eigenvalues of $\mathbf{A}$, and $\mathbf{U}$ is unitary.

   We define a matrix $\mathbf{B} = \mathbf{U}\widetilde{\mathbf{T}}\mathbf{U}^*$, where $\widetilde{\mathbf{T}} = \mathbf{T} - \text{diag}(\epsilon_1, \ldots, \epsilon_n)$. Here, ϵ_i are chosen so that

   (a) $\epsilon_i > 0$ and $\epsilon_i \le \epsilon$, $1 \le i \le n$; and
   (b) $\lambda_i - \epsilon_i \ne \lambda_j - \epsilon_j$ if $i \ne j$, where λ_i and λ_j are the i^{th} and j^{th} eigenvalues of $\mathbf{A}$, respectively.

Thus, $\mathbf{B}$ has distinct eigenvalues by construction, which implies that $\mathbf{B}$ is diagonalizable.

We have that $\mathbf{A} - \mathbf{B} = \mathbf{U}\mathbf{D}\mathbf{U}^*$, where $\mathbf{D} = \text{diag}\,(\epsilon_1, \ldots, \epsilon_n)$. Thus, $(\mathbf{A} - \mathbf{B})$ is Hermitian, and thus, normal, which implies that the singular values of $(\mathbf{A} - \mathbf{B})$ are the ϵ_i's. From this, we see that

$$\begin{aligned} ||\mathbf{A} - \mathbf{B}||_2 &= \sigma_{\max}\,(\mathbf{A} - \mathbf{B}) \\ &= |\lambda_{\max}\,(\mathbf{A} - \mathbf{B})| \\ &= \max_{1 \leq i \leq n} \epsilon_i \\ &\leq \epsilon, \end{aligned}$$

by construction of the ϵ_i's.