System Theory, KOM5108
Lecture #2

Instructor: Dr. Yavuz Eren

Lecture Book: **"Linear System Theory, W.J. Rugh''**

Spring 2024

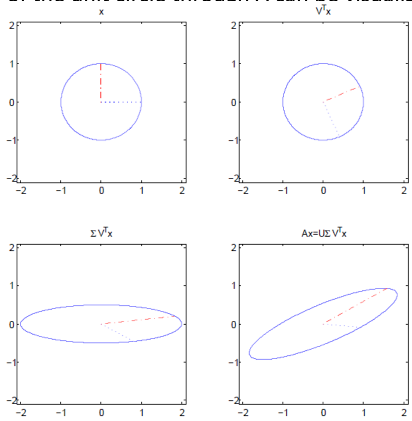
Singular Value Decomposition

Theorem: Given A $\in \mathcal{F}^{m\times n}$, there exists unitary matrices $U = [u_1 \quad u_2 \quad \cdots \quad u_m] \in \mathcal{F}^{m\times n}$ and $V = [v_1 \quad v_2 \quad \cdots \quad v_m] \in \mathcal{F}^{n\times n}$ and a matrix Σ, where

$$\Sigma \triangleq \begin{bmatrix} \Sigma_1 & 0 \\ 0 & 0 \end{bmatrix}, \Sigma_1 = diag[\sigma_1, \sigma_2, \dots, \sigma_k], k \triangleq rank(A)$$

and $\sigma_1 \geq \sigma_2 \geq \cdots \sigma_k > 0$ such that $A = U\Sigma V^*$.

Consider the action of a matrix A on vectors forming the unit circle. If A is decomposed as $A = U\Sigma V^T$, the transformation of the unit circle through A can be visualized as follows:



Examples :

Hermitian Matrices
A matrix $A \in \mathbb{C}^{n\times n}$ is said to be Hermitian if $A = A^*$.

Teorem: All eigenvalues of a Hermitian matrix are real.
Proof: Let us to take arbitrary $A = A^* \in \mathbb{C}^{n\times n}$ and consider its eigenvalue-eigenvector equation $Ax = \lambda x$. Then, by premultiplying by x^* we get $x^*Ax = \lambda x^*x$. Now, take the complex conjugate transposes of both sides to obtain $(x^*Ax)^* = (\lambda x^*x)^*$. But $(x^*Ax)^* = x^*A^Tx = \lambda x^*x$ since $A = A^*$. Furthermore, since $(\lambda x^*x)^* = \lambda x^*x$. Since x is nonzero, we get $\lambda = \lambda$.

By using the unitary matrices($U \in \mathbb{C}^{n\times n}$; $U^*U = UU^* = I$), a useful fact about the spectral decomposition of Hermitian matrices can be given as follows:

Theorem: Any Hermitian matrix $A = A^* \in \mathbb{C}^{n\times n}$ can be spectrally decomposed as $A = U\Lambda U^*$ where U is unitary and $\Lambda = diag(\lambda_1, \lambda_2, \dots, \lambda_n)$
Proof: (Proof by induction over the dimension n). For a Hermitian matrix $\mathbb{C}^{1\times 1}$, the result is obvious. Because if $A \in \mathbb{C}^{1\times 1}$ is hermitian it is a real scalar. Then, Λ is equal to itself and $U = 1$, which is unitary.
Now assume that the statement is true for all Hermitian matrices in $\mathbb{C}^{(n-1)\times(n-1)}$ and also consider $A = A^* \in \mathbb{C}^{n\times n}$. Let λ_1 be an eigenvalue of A and x_1 eigenvector associated with λ_1. By the above theorem, λ_1 is real and x_1 can be chosen as $x_1^*x_1 = 1$ without loss of generality. Now let X be unitary matrix with x_1 as its first column

$$X \triangleq [x_1 \quad x_1 \quad \cdots \quad x_n] \in \mathbb{C}^{n\times n}$$

Then, the first column of the product X^*AX gives

$$X^* A x_1 = X^*(\lambda_1 x_1) = \lambda_1 X^* x_1 = \lambda_1 e_1$$

where e_1 denotes the first column of I_n. Moreover, the first row of X^*AX is

$$x_1^* A X = (Ax_1)^* X = \lambda_1 x_1^* X = \lambda_1 e_1^*$$

since $A = A^*$. We then have

$$X^* A X = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \tilde{A} \end{bmatrix}$$

where $\tilde{A} = \tilde{A}^* \in \mathbb{C}^{(n-1)\times(n-1)}$. By the induction, we can state $\tilde{A}$ as $\tilde{A} = \widetilde{U}\widetilde{\Lambda}\widetilde{U}^*$ where $\widetilde{U}$ is unitary and $\widetilde{\Lambda}$ is real diagonal. Then we obtain the following form

$$A = X \underbrace{\begin{bmatrix} I & 0 \\ 0 & \tilde{U} \end{bmatrix}}_{U} \underbrace{\begin{bmatrix} \lambda_1 & 0 \\ 0 & \tilde{\Lambda} \end{bmatrix}}_{\Lambda} \underbrace{\begin{bmatrix} I & 0 \\ 0 & \tilde{U} \end{bmatrix}^* X^*}_{U^*}.$$

Note that, U is unitary and Λ is is real diagonal so that the diagonal entries are the eigenvalues of A.(A is Hermitian.)

Congruence Transformation

Two matrices $\mathrm{A}, \mathrm{B} \in \mathbb{C}^{n\times n}$ are said to be congruent to each other if there exists a nonsingular $\mathrm{T} \in \mathbb{C}^{n\times n}$ such that $B = T^*AT$. The transformation $A \to T^*AT$ is called a congruence transformation of A under T.

Example:

Note that, all eigenvalues of a Hermitian matrix are real, we can group the ones that are positive, negative and zero. This brings the definition of the inertia of a Hermitian matrix.

Definition(Inertia): Given a matrix $\mathrm{A} = A^* \in \mathbb{C}^{n\times n}$, the inertia of A is the triplet $in(A) = (n_+, n_-, n_0)$ where n_+, n_- and n_0 denote the number of positive, negative and zero eigenvalues, respectively.

Note that, the inertia of a symmetric matrix remains unchanged under congruence transformation:

Lemma: Given two Hermitian matrices $\mathrm{A}, \mathrm{B} \in \mathbb{C}^{n\times n}$: $A \text{ and } B \text{ congruent} \Leftrightarrow in(A) = in(B)$.

Proof: (HW)

Sign Definiteness
A matrix $A = A^* \in \mathbb{C}^{n\times n}$ is said to be

(i) **Positive definite** *if*
$$x^T A x > 0 \quad \forall x \in \mathbb{C}^n, \quad x \neq 0$$
(ii) **Negative definite** *if*
$$x^T A x < 0 \quad \forall x \in \mathbb{C}^n, \quad x \neq 0$$
(iii) **Positive semidefinite** *if*
$$x^T A x \geq 0 \quad \forall x \in \mathbb{C}^n, \quad x \neq 0$$
(iv) **Negative semidefinite** *if*
$$x^T A x \leq 0 \quad \forall x \in \mathbb{C}^n, \quad x \neq 0$$

By the definition of Rayleight-Ritz inequality, the sign-definiteness of a Hermitian matrix is determined completely by the signs of its eigenvalues:

Lemma: A matrix $A = A^* \in \mathbb{C}^{n\times n}$ is

(i) Positive definite if and only if $\lambda_i > 0$ for all $i = 1:n$

(ii) Negative definite if and only if $\lambda_i < 0$ for all $i = 1:n$

(iii) Positive semi-definite if and only if $\lambda_i \geq 0$ for all $i = 1:n$

(iv) Negative semi-definite if and only if $\lambda_i \leq 0$ for all $i = 1:n$

where λ_i denote the eigenvalue of A

Corollary : The matrices $A = A^* \in \mathbb{C}^{n\times n}$ and $B = B^* \in \mathbb{C}^{m\times m}$,

$$\begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix} \geq 0 \quad \Longleftrightarrow \quad \begin{cases} A \geq 0 \\ B \geq 0 \end{cases}$$

Lemma: Let $P = P^* > 0$. Then the matrix T^*PT is also symmetric positive definite for all nonsingular $T \in \mathbb{C}^{n\times n}$.
Proof: If $P > 0$, then $in(P) = \{n, 0,0\}$. Since inertia of a Hermitian matrix is invariant under congruence transformation, $in(T^*PT) = in(P)$, hence $T^*PT \geq 0$.

Similar to the positive(or non-negative) scalars, we can define the square root of a positive semidefinite matrix:

Definition(Matrix Square Root): The square root of $P = P^* \geq 0$ is defined as the matrix $P^{1/2} \triangleq U^T \Lambda^{1/2} U \geq 0$ where $P = U\Lambda U^*$ is the spectral decomposition of P and
$\Lambda^{1/2} \triangleq diag\{\sqrt{\lambda_1}, \dots, \sqrt{\lambda_n}\}$.

Lemma (Schur Complement Formula): Given $A = A^* \in \mathbb{C}^{n\times n}$, $B \in \mathbb{C}^{n\times m}$ and $C \in \mathbb{C}^{m\times m}$, the following statements are equivalent:

$i)\begin{bmatrix} A & B \\ B^* & C \end{bmatrix} > 0,$

$ii) A > 0 \;\&\; C - B^*A^{-1}B > 0,$

$iii) C > 0 \;\&\; A - BC^{-1}B^* > 0.$

Proof: $(i) \Leftrightarrow (ii)$: By using congruence transformation

$$T^*\begin{bmatrix} A & B \\ B^* & C \end{bmatrix}T = \begin{bmatrix} A & 0 \\ 0 & C - B^*A^{-1}B \end{bmatrix}, \quad \text{where} \quad T := \begin{bmatrix} I & -A^{-1}B \\ 0 & I \end{bmatrix}$$

$(i) \Leftrightarrow (iii)$:

$$J^*\begin{bmatrix} A & B \\ B^* & C \end{bmatrix}J = \begin{bmatrix} C & B^* \\ B & A \end{bmatrix}, \quad \text{where} \quad J := \begin{bmatrix} 0 & I \\ I & 0 \end{bmatrix}$$

and repeating the proof of $(i) \Leftrightarrow (ii)$.

Theorem: $P \geq 0 \;\Leftrightarrow Trace(PS) \geq 0 \; for\; every\; S \geq 0.$

Proof: (if) Assume $P \geq 0$. Let $S \geq 0$ be given and decompose it as $S = U\Lambda U^T = \sum_{i=1}^{n} \lambda_i u_i u_i^T$. Then, $0 \leq u_i^T P u_i = Trace(P u_i u_i^T)$ for each i. Since each $\lambda_i \geq 0$, we get

$$0 \leq \sum \lambda_i \,\mathbf{Trace}(P u_i u_i^T) = \mathbf{Trace}\left(P \sum_i \lambda_i u_i u_i^T\right) = \mathbf{Trace}(PS).$$

(only if) Assume $Trace(PS) \geq 0 \; for\; every\; S \geq 0$. Since $xx^T \geq 0$ for each $for\; every\; x \in \mathbb{R}^n$, we have

$$0 \leq \mathbf{Trace}(Pxx^T) = x^T P x \quad \forall x \in \mathbb{R}^n.$$

Hence, P$\geq$ 0 by definition.

Matrix Norms

Matrix norm is used as a way of measuring the size of a matrix. A marix norm is any function $\|.\|$: $\mathbb{C}^{n\times n} \to \mathbb{R}$ that satisfy the following properties:

(i) $\|A\| \geq 0$ *for all* $A \in \mathbb{C}^{n\times n}$

(i') $\|A\| = 0$ *if and only if* $A = 0$

(ii) $\|\alpha A\| = |\alpha|\, \|A\|$ *for all* $A \in \mathbb{C}^{n\times n}$ *and* $\alpha \in \mathbb{C}$

(iii) $\|A + B\| \leq \|A\| + \|B\|$ *for all* $A, B \in \mathbb{C}^{n\times n}$

(iv) $\|AB\| \leq \|A\|\, \|B\|$ *for all* $A, B \in \mathbb{C}^{n\times n}$.

Some examples of matrix norms:
1) 1-norm:

$$\|A\|_1 := \sum_{i=1}^{n}\sum_{j=1}^{n} |a_{ij}|.$$

This norm satisfies the first three conditions without needing detailed proof, for the last condition:

$$\|AB\|_1 = \sum_{i,j}\left|\sum_k a_{ik}b_{kj}\right| \leq \sum_{i,j,k} |a_{ik}b_{kj}| \leq \sum_{i,j,k,l} |a_{ik}b_{lj}|$$
$$= \left(\sum_{i,k} |a_{ik}|\right)\left(\sum_{l,j} |b_{lj}|\right) = \|A\|_1\,\|B\|_1,$$

where the first inequality is obtained by the scalar triangular inequality, and the second one by adding positive terms to the sum.

2) 2-norm or Frobenius norm($\|A\|_F$): $$\|A\|_2 := \left(\sum_{i=1}^{n}\sum_{j=1}^{n} |a_{ij}|^2\right)^{1/2}.$$

Note that $\|A\|_F^2 = Trace(A^*A)$.

Frobenius norm is also a valid matrix norm which can be verified as follows:
(Similar to Cauchy-Schwarz Ieequality)

$$\|AB\|_F^2 = \sum_{i,j} \left| \sum_{k=1}^{n} a_{ik} b_{kj} \right|^2 \leq \sum_{i,j} \left(\sum_k |a_{ik}|^2 \right) \left(\sum_m |b_{mj}|^2 \right)$$
$$= \left(\sum_{i,k} |a_{ik}|^2 \right) \left(\sum_{m,j} |b_{mj}|^2 \right) = \|A\|_F^2 \|B\|_F^2.$$

3) İnfinity-Norm($\|A\|_\infty \triangleq max_{i,j}|a_{i,j}|$) is a vector norm on $\mathbb{C}^{n\times n}$, but not a matrix norm. This is because the submultiplicative condition $\|AB\|_\infty \leq \|A\|_\infty \|B\|_\infty$ is violated(For instance,$A = B = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$)

However, a modified $\infty - norm$ can be defined as:

$$\|A\|_{\sim\infty} := n\|A\|_\infty = n \max_{1\leq i,j\leq n} |a_{ij}|,$$

Then, $\|A\|_{\sim\infty}$ is an appropriate matrix norm as prove as follows:

$$\|AB\|_{\sim\infty} = n \max_{i,j} \left| \sum_k a_{ik} b_{kj} \right| \leq n \max_{i,j} \sum_k |a_{ik} b_{kj}|$$
$$\leq n \max_{i,j} \sum_k \|A\|_\infty \|B\|_\infty$$
$$= n\|A\|_\infty \, n\|B\|_\infty = \|A\|_{\sim\infty} \|B\|_{\sim\infty}$$

Induced Matrix Norm

Let $\|.\|_*$ be a vector norm on $\mathbb{C}^n$. Then, the induced norm on $\mathbb{C}^{n\times n}$ is defined as

$$\|A\|_{i*} := \max_{x\neq 0} \frac{\|Ax\|_*}{\|x\|_*}.$$

Moreover, it can be shown that:

$$\|A\|_{i*} = \max_{\|x\|_*\leq 1} \|Ax\|_* = \max_{\|x\|_*=1} \|Ax\|_*.$$

Note that, it is easy to Show that induced matrix norms are legitimate matrix norms as defined the above definition(required conditions for the norm).

Condition 1 is clearly satisfied since the induced norm is always nonnegative by definition and $Ax = 0$ for all nonzero $x \in \mathbb{C}^n$ iff $A = 0$.

Condition 2 is also satisfied since

$$\|\alpha A\|_{i*} = \max_{\|x\|_*=1} \|\alpha Ax\|_* = \max_{\|x\|_*=1} |\alpha| \, \|Ax\|_* = |\alpha| \max_{\|x\|_*=1} \|Ax\|_* = |\alpha| \, \|A\|_{i*}.$$

Condition 3 is also satisfied as follows:

$$\|A+B\|_{i*} = \max_{\|x\|_*=1} \|(A+B)x\|_* \leq \max_{\|x\|_*=1} (\|Ax\|_* + \|Bx\|_*)$$
$$\leq \max_{\|x\|_*=1} \|Ax\|_* + \max_{\|x\|_*=1} \|Bx\|_* = \|A\|_{i*} + \|B\|_{i*}.$$

Condition 4(submultiplicavity) can be demonstrated as

$$\|AB\|_{i*} = \max_{x\neq 0} \frac{\|ABx\|_*}{\|x\|_*} = \max_{x\neq 0} \frac{\|ABx\|_*}{\|Bx\|_*}\frac{\|Bx\|_*}{\|x\|_*} \leq \max_{y\neq 0} \frac{\|Ay\|_*}{\|y\|_*} \max_{x\neq 0} \frac{\|Bx\|_*}{\|x\|_*} = \|A\|_{i*}\|B\|_{i*}.$$

It also can be stated as by the definition of induced norm

$$\|Ax\|_* \leq \|A\|_{i*}\|x\|_* \quad \forall x \in \mathbb{C}^n.$$

Most common-used induced norms are the p-norms: (p=1,2,inf)

$$\|A\|_{i1} := \max_{x\neq 0} \frac{\|Ax\|_1}{\|x\|_1}, \quad \|A\|_{i2} := \max_{x\neq 0} \frac{\|Ax\|_2}{\|x\|_2}, \quad \text{and} \quad \|A\|_{i\infty} := \max_{x\neq 0} \frac{\|Ax\|_\infty}{\|x\|_\infty}.$$

Examples : Simple characterizations for p-norms:

$$\|A\|_{i1} = \max_{j=1:n} \sum_{i=1}^{n} |a_{ij}|.$$

Proposition: The induced p-norms can be calculated as follows:

(i) The induced 1-norm is the maximum column sum:

$$\|A\|_{i1} = \max_{j=1:n} \sum_{i=1}^{n} |a_{ij}|$$

(ii) The induced 2-norm(spectral norm) is the square root of the maximum eigenvalue of A^*A:

$$\|A\|_{i2} = \sqrt{\lambda_{max}(A^*A)}$$

(iii) The induced inf-norm is the maximum row-sum:

$$\|A\|_{i\infty} = \max_{i=1:n} \sum_{j=1}^{n} |a_{ij}|$$

Examples:

Ellipsoids: Given a matrix $P \in \mathbb{R}^{n\times n}$, where $P = P^T \geq 0$ the set

$$\mathcal{E}_P \triangleq \{x \in \mathbb{R}^n : x^T P x \leq 1\}$$

is called as ellipsoid associated with P.

Not that, the ellipsoid defined abov is centered at the origin of $\mathbb{R}^n$. The ellipsoid centered at different points as given as

$$\mathcal{E}_{P,x_c} \triangleq \{x \in \mathbb{R}^n : (x - x_c)^T P (x - x_c) \leq 1\}.$$

If $P > 0$, the volume of an ellipsoid is $vol(\mathcal{E}_P) = \alpha_n (detP^{-1})^{1/2}$ where α_n is the volume of the n-dimensional unit 2-ball, i. e. ,

$$\alpha_n = \begin{cases} \dfrac{\pi^{n/2}}{(n/2)!} & \text{if } n \text{ is even} \\[2ex] \dfrac{2^n \pi^{(n-1)/2}((n-1)/2)!}{n!} & \text{if } n \text{ is odd.} \end{cases}$$

Proposition: Given two matrices $A \geq 0$ and $B \geq 0$, $\mathcal{E}_A \subseteq \mathcal{E}_B \Leftrightarrow B \leq A$.

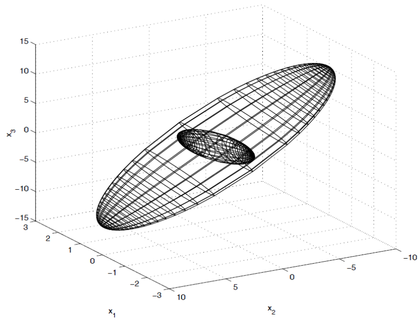
Proof: (only if) Suppose $B \leq A$ and let x $\subseteq \mathcal{E}_A$. Then, $x^T B x \leq x^T A x \leq 1$. Therefore, x $\subseteq \mathcal{E}_B$, so that $\mathcal{E}_A \subseteq \mathcal{E}_B$.

(if) Suppose $\mathcal{E}_A \subseteq \mathcal{E}_B$ but that $B \nleq A$, i.e., there exists nonzero x_* such that $x_*^T B x_* > x_*^T A x_*$. Now scale x_* to $\tilde{x}_*$ so that(It is assumed $Ax_* \neq 0$) $\tilde{x}_*^T A \tilde{x}_*$=1. Then, $\tilde{x}_* \in \mathcal{E}_A$ but , $\tilde{x}_* \notin \mathcal{E}_B$ since $\tilde{x}_*^T B \tilde{x}_*$> $\tilde{x}_*^T A \tilde{x}_*$=1. This is a conrtradiction due to the assumption $\mathcal{E}_A \subseteq \mathcal{E}_B$.

Ex: Consider the matrices

$$A = \begin{bmatrix} 1.3244 & 1.1302 & 0.9364 \\ 1.1302 & 1.6114 & 0.8643 \\ 0.9364 & 0.8643 & 0.8834 \end{bmatrix} \quad B = \begin{bmatrix} 1.0400 & 1.0502 & 0.7666 \\ 1.0502 & 1.3019 & 0.9304 \\ 0.7666 & 0.9304 & 0.6726 \end{bmatrix}$$

By calculation, we know that $B < A$, so that $\mathcal{E}_A \subseteq \mathcal{E}_B$ and the inclusion ca n be illustrated as follows:



Remark: When $P > 0$ the definition can be thought of as the unit ball of the vector norm on $\mathbb{C}^n$ as follows: $\|x\| \triangleq (x^*Px)^{1/2}$.

Note that, when $P \geq 0$ then one can find nonzero vectors x $\in \mathbb{C}^n$ such that $\|x\| = 0$, violating condition for being a norm. It is then called as semi-norm on $\mathbb{C}^n$.

Lemma: Given $P = P^T = \begin{bmatrix} P_1 & P_2 \\ P_2^T & P_3 \end{bmatrix} > 0$, let $\mathcal{E} \triangleq \{x: x^TPx \leq 1\}$. Then, the projection of $\mathcal{E}$ onto the $x_1 - subspace$ is given by

$$\mathcal{E}_{x_1} = \{z \in \mathbb{R}^n: z^T(P_1 - P_2P_3^{-1}P_2^T)z \leq 1\}$$

Proof: the projection on $\mathcal{E}$ onto x_1 subspace is given by

$$\begin{aligned} \min_{Q} \quad & \mathbf{vol}\{z : z^TQz \leq 1\} \\ s.t. \quad & x \in \mathcal{E} \Rightarrow x_1^TQx_1 \leq 1. \end{aligned}$$

So the last condition is equivalent to

$$\{x : x^TPx \leq 1\} \subseteq \left\{x : x^T \begin{bmatrix} Q & 0 \\ 0 & 0 \end{bmatrix} x \leq 1\right\}$$

Or by the above Proposition:

$$\begin{bmatrix} Q & 0 \\ 0 & 0 \end{bmatrix} \leq \begin{bmatrix} P_1 & P_2 \\ P_2^T & P_3 \end{bmatrix} \Longleftrightarrow \begin{bmatrix} Q - P_1 & -P_2 \\ -P_2^T & -P_3 \end{bmatrix} < 0.$$

By the schur complement formula, this is equivalent to

$$Q - (P_1 - P_2 P_3^{-1} P_2^T) \leq 0.$$

And, the minumum volüme is $Q = P_1 - P_2 P_3^{-1} P_2^T$.

Ex. :