Modern Control Systems

Matthew M. Peet
Arizona State University

Lecture 4: Back to Matrices

Eigenvalues and Eigenvectors

Eigenvalues and Spectrum are very different for matrices vs. operators.

Definition 1.

For a matrix $A \in \mathbb{R}^{n \times n}$, $\lambda \in \mathbb{R}$ is an **eigenvalue** of A, denoted $\lambda \in \text{eig}(A)$ if there exists some **eigenvector** $v_\lambda \in \mathbb{R}^n$ such that

$$Av_\lambda = \lambda v_\lambda$$

Alternative Representations:

- $\lambda \in \text{eig}(A)$ if $\text{Ker}(\lambda I - A) \neq \{0\}$
- For matrices, $\lambda \in \text{eig}(A)$ if $\det(\lambda I - A) = 0$

Recall:

- $\det(A) = \prod \lambda_i(A)$
- $\text{trace}(A) = \sum \lambda_i(A)$

Eigenvalues and Eigenvectors

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}$$

```
[U,V]=eigs([1 2 3; 4 5 6; 7 8 9])
```

$$U = \begin{bmatrix} -.23 & -.78 & .41 \\ -.53 & -.09 & -.82 \\ -.82 & .61 & .41 \end{bmatrix}; \qquad V = \begin{bmatrix} 16.1 & 0 & 0 \\ 0 & -1.11 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

which means:

$$\lambda_1(A) = 16.1, \quad \lambda_2(A) = -1.11, \quad \lambda_3(A) = 0,$$

$$v_1 = \begin{bmatrix} -.23 \\ -.53 \\ -.82 \end{bmatrix}, \quad v_2 = \begin{bmatrix} -.78 \\ -.09 \\ .61 \end{bmatrix}, \quad v_3 = \begin{bmatrix} .41 \\ -.82 \\ .41 \end{bmatrix}$$

Characteristic Polynomials

Definition 2.

The **characteristic polynomial** of a matrix A is denoted

$$\text{char}A(\lambda) := \det(\lambda I - A)$$

Theorem 3.

$\lambda \in \text{eig}(A)$ if and only if $\text{char}A(\lambda) = 0$

Proof.

- By definition, $\lambda \in \text{eig}(A)$ means that $\exists v \neq 0 \in \mathbb{R}^n$ such that $Av = \lambda I$.
- $(\lambda I - A)v = 0$ for some $v \neq 0$ if and only if 0 is an eigenvalue of $\lambda I - A$.
- If 0 is an eigenvalue of $\lambda I - A$, then $\text{char}A(\lambda) = \det(\lambda I - A) = \prod_{\gamma_i \in \text{eig}(\lambda I - A)} \gamma_i = 0$.
- Likewise, if $\text{char}A(\lambda) = \det(\lambda I - A) = 0$, then $\lambda I - A$ has at least one zero eigenvalue.

□

Characteristic Polynomials

Eigenvectors are **NOT** unique

- If v is an eigenvector, then so is αv
- If v_1 and v_2 are both eigenvectors for λ, so is $\alpha v_1 + \beta v_2$.

To avoid the problem of uniqueness, we use eigenspace

Definition 4.

For any $\lambda \in \text{eig}(A)$, there is a unique subspace, S_λ, called an **eigenspace**, such that $x \in S_\lambda$ if and only if

$$Ax = \lambda x.$$

Theorem 5.

If $\cup_i \{S_i\} = \mathbb{R}^n$, then there exist eigenvectors v_i such that

$$V = [v_1, \cdots, v_n] \qquad \textit{is invertible.}$$

Diagonalizability

Then

$$AV = [Av_1, \cdots, Av_n] = [\lambda_1 v_1, \cdots, \lambda_n v_n] = V\Lambda, \quad \text{where} \quad \Lambda = \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix}.$$

If V is invertible, then

$$V^{-1}AV = \Lambda$$

Thus if the eigenspaces span the space, the matrix can be diagonalized via a *Similarity Transformation*.

- But when does this happen?
- Not all matrices can be diagonalized.

Example:

$$A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$$

$$\lambda_1 = 0, \qquad v_1 = \begin{bmatrix} -1 \\ 0 \end{bmatrix}, \qquad \lambda_2 = 0, \qquad v_2 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

Jordan Form

Diagonalizibility is an important tool.

- A change of bases
- In new basis, the operator/matrix multiplies coordinate by a simple factor.
- Can all operators be so simple?
  - **NO**

An alternative which applies to ALL matrices is the Jordan Form

Definition 6.

A **Jordan Block** has the form

$$\lambda I + N = \begin{bmatrix} \lambda & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ & & & \lambda \end{bmatrix}$$

Almost Diagonal.

Jordan Form

$$N = \begin{bmatrix} 0 & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ & & & 0 \end{bmatrix}$$

is called **Nilpotent** as

- $N^d = 0$ for some $d > 0$.

Example

$$N = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \; N^2 = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \; N^3 = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \; N^4 = 0$$

Jordan Form

λ is the sole eigenvalue of $\lambda I + N$ with multiplicity n and eigenvectors

$$v_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \quad v_2 = \begin{bmatrix} -1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \quad v_3 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \quad v_4 = \begin{bmatrix} -1 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

or a 1-Dimensional eigenspace $S_\lambda := \text{span}\,\{e_1\}$.

- Jordan blocks capture the part of a matrix which cannot be diagonalized.

Theorem 7.

For any $A \in \mathbb{C}^{n \times n}$, there exists an invertible T and Jordan Blocks J_i such that

$$TAT^{-1} = J = \begin{bmatrix} J_1 & & \\ & \ddots & \\ & & J_p \end{bmatrix}$$

ANY matrix is Jordan-Block Diagonalizable.
Different Jordan Blocks may have the same eigenvalue.

Symmetric Matrices

Definition 8.

A Matrix, $A \in \mathbb{R}^{n\times n}$ ($A \in \mathbb{C}^{n\times n}$) is self-adjoint, denoted $A \in \mathbb{S}^n$ ($A \in \mathbb{H}^n$) if $A = A^T$ ($A = A^*$).

- Real self-adjoint matrices are called symmetric. ($\mathbb{S}$)
- Complex self-adjoint matrices are called Hermetian. ($\mathbb{H}$)

Lemma 9.

Both Hermetian and Symmetric Matrices have **real** *Eigenvalues.*

Proof.

We show that is $A \in \mathbb{H}$ and $Av = \lambda v$, then $\lambda = \lambda^*$.

- If $Av = \lambda v$, then

$$\lambda v^* v = v^* A v = (A^* v)^* v = (Av)^* v = (\lambda v)^* v = \lambda^* v^* v$$

- Hence $\lambda = \lambda^*$, which means λ is real.

□

Unitary Matrices

Unitary matrices are a special case of coordinate transformation.

Definition 10.

Two vectors, $x, y \in R^n$ are

- **orthogonal** if $x^T y = 0$
- **orthonormal** if they are orthogonal and $\|x\| = \|y\| = 1$

A basis, $\{v_i\}$ is an **orthonormal basis** if all v_i are orthonormal.

- A coordinate transformation to an orthonormal basis is a *unitary* operator.

Given a basis $\{v_1, \cdots, v_n\}$, we can construct an orthonormal basis

$$
\begin{aligned}
x_1 &= \frac{v_1}{\|v_1\|} \\
x_2 &= \frac{v_2 - (v_2^T x_1)x_1}{\|v_2 - (v_2^T x_1)x_1\|} \\
x_3 &= \frac{v_3 - (v_3^T x_2)x_2 - (v_3^T x_1)x_1}{\|v_3 - (v_3^T x_2)x_2 - (v_3^T x_1)x_1\|}
\end{aligned}
$$

- Clearly, all vectors are of unit length
- To see orthogonality, note

$$
x_2^T x_1 = \frac{v_2^T x_1 - (v_2^T x_1)x_1^T x_1}{\|v_2 - (v_2^T x_1)x_1\|} = \frac{v_2^T x_1 - v_2^T x_1}{\|(v_2 - v_2^T x_1)x_1\|} = \frac{0}{v_2 - \|(v_2^T x_1)x_1\|}
$$

Unitary Matrices

Definition 11.

A matrix U is **Unitary** if $U^*U = I$.

- The columns of U form an orthonormal basis

$$\begin{bmatrix} u_1^* \\ u_2^* \\ \vdots \\ u_n^* \end{bmatrix} \begin{bmatrix} u_1 & u_2 & \cdots & u_n \end{bmatrix} = \begin{bmatrix} u_1^*u_1 & u_1^*u_2 & \cdots & u_1^*u_n \\ u_2^*u_1 & u_2^*u_1 & & u_2^*u_n \\ \vdots & & \ddots & u_1^*u_n \\ u_n^*u_1 & & \cdots & u_n^*u_n \end{bmatrix} = \begin{bmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{bmatrix}$$

- $U^* = U^{-1}$
  - The columns of U^* also form an orthonormal basis

Length Preservation: A basis change which preserves norms.

- If $y = Ux$, then

$$\|y\|^2 = x^*U^*Ux = x^*x = \|x\|^2$$

- Why the $\|\cdot\|_2$-norm?

Symmetric matrices can be diagonalized via unitary matrices.

Theorem 12.

If $A \in \mathbb{H}$, then there exists a unitary matrix U and a real diagonal Λ such that

$$A = U \Lambda U^*$$

There is also a spectral theorem for operators.

- What is a diagonal operator?

Singular Value Decomposition

Theorem 13.

Let $A \in \mathbb{R}^{m,n}$ and $p = \min(m,n)$. There exist unitary matrices $U \in \mathbb{R}^{m \times m}$ and $V \in \mathbb{R}^{n \times n}$ such that $A = U\Sigma V^T$, where

$$\Sigma_{i,j} = \begin{cases} 0 & i \neq j \\ \sigma_i > 0 & i = j \end{cases} \qquad \Sigma = \begin{bmatrix} \sigma_1 & & & 0 & \cdots & 0 \\ & \ddots & & & & \vdots \\ & & \sigma_p & 0 & \cdots & 0 \end{bmatrix}$$

Assume that $\sigma_1 > \sigma_2 > \cdots > \sigma_p$.

Singular Value Decomposition

Proof.

By Spectral Theorem,

$$A^T A = V \begin{bmatrix} \Lambda & \\ & 0 \end{bmatrix} V^T$$

where $\Lambda > 0$ is diagonal and V is unitary. Let $\Sigma = \Lambda^{\frac{1}{2}}$. Then

$$V^T A^T A V = \begin{bmatrix} \Sigma^2 & \\ & 0 \end{bmatrix}$$

Let $X = AV \begin{bmatrix} \Sigma^{-1} & \\ & I \end{bmatrix}$. Then

$$\begin{aligned} x^T X &= \begin{bmatrix} \Sigma^{-1} & \\ & I \end{bmatrix}^T V^T A A V \begin{bmatrix} \Sigma^{-1} & \\ & I \end{bmatrix} \\ &= \begin{bmatrix} \Sigma^{-1} & \\ & I \end{bmatrix}^T \begin{bmatrix} \Sigma^2 & \\ & 0 \end{bmatrix} \begin{bmatrix} \Sigma^{-1} & \\ & I \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & I \end{bmatrix}. \end{aligned}$$

□

Singular Value Decomposition

Proof.

Thus if $X = \begin{bmatrix} x_1 & \cdots x_n \end{bmatrix}$, then

$$x_i^T x_j = \begin{cases} 1 & i = j, \;\; i = 1 \dots k < n \\ 0 & \text{otherwise} \end{cases}$$

Thus the first k columns are orthonormal and the rest are zero. Now define $U_1 = \begin{bmatrix} x_1 & \cdots & x_k \end{bmatrix}$ so that $X = \begin{bmatrix} U_1 & 0 \end{bmatrix}$. Now complete the basis as $U = \begin{bmatrix} U_1 & U_2 \end{bmatrix}$, where $U_2 = \begin{bmatrix} v_{k+1} & \cdots & v_n \end{bmatrix}$ is a arbitrarily chosen set of orthonormal vectors. Then

$$X = AV \begin{bmatrix} \Sigma^{-1} & \\ & I \end{bmatrix} = \begin{bmatrix} U_1 & 0 \end{bmatrix}$$

$$\begin{aligned} A &= \begin{bmatrix} U_1 & 0 \end{bmatrix} \begin{bmatrix} \Sigma & \\ & I \end{bmatrix} V^T = \begin{bmatrix} U_1 & 0 \end{bmatrix} \begin{bmatrix} \Sigma & \\ & 0 \end{bmatrix} V^T \\ &= \begin{bmatrix} U_1 & U_2 \end{bmatrix} \begin{bmatrix} \Sigma & \\ & 0 \end{bmatrix} V^T = U \begin{bmatrix} \Sigma & \\ & 0 \end{bmatrix} V^T \end{aligned}$$

The Maximum Singular Value

The σ_i^2 are the eigenvalues of $A^T A$ or AA^T.

Definition 14.

We denote the Maximum Singular Value of a Matrix, M, as

$$\bar{\sigma}(M) = \max_i \sigma_i(M)$$

The maximum singular value of a matrix is a matrix norm with many pleasing properties.

- An induced norm

$$\bar{\sigma}(A) = \sup_v \frac{\|Av\|}{\|v\|} = \sup_{\|v\|=1} \|Av\|$$

Matrix Norms

The set of matrices is a normed space with a variety of norms.

- Induced Norm (from a vector norm)

$$\|M\|_{Xi} = \max_{v\in\mathbb{R}^n} \frac{\|Mv\|_X}{\|v\|_X}$$

  - Maximum Singular Value Norm is induced from the Euclidean norm $\|x\|_2 = \sqrt{x^T x}$

$$\bar{\sigma}(M) = \|M\|_{2i} = \max_{v\in\mathbb{R}^n} \frac{\|Mv\|_2}{\|v\|_2}$$

  - $\|M\|_{1i}$ is the Maximum Column Sum Norm

$$\|M\|_{1i} = \max_{v\in\mathbb{R}^n} \frac{\|Mv\|_1}{\|v\|_1} = \max_j \sum_i |M_{i,j}|$$

  - $\|M\|_{\infty i}$ is the Maximum Column Sum Norm

$$\|M\|_{\infty i} = \max_{v\in\mathbb{R}^n} \frac{\|Mv\|_\infty}{\|v\|_\infty} = \max_i \sum_j |M_{i,j}|$$

Other Matrix Norms

- Frobenius Norm

$$\|M\|_F = \|M\|_2 = \sqrt{\sum_i \sum_j |M_{i,j}|^2} = \sqrt{\text{trace}(M^* M)} = \sqrt{\sum_i \sigma_i^2(M)}$$

- Taxicab Norm

$$\|M\|_1 = \sum_i \sum_j |M_{i,j}|$$

- Maximum Value Norm

$$\|M\|_\infty = \max_{i,j} |M_{i,j}|$$

- p-norm

$$\|M\|_p = \sqrt[p]{\sum_i \sum_j |M_{i,j}|^p}$$

- Spectral Radius (False Norm)

$$\rho(M) = \max_{\lambda \in \text{eig}(A)} |\lambda|$$

However, $\rho(M) \le \|M\|$ for any matrix norm.

Positive Matrices

Definition 15.

A symmetric matrix $P \in \mathbb{S}^n$ is **Positive Semidefinite**, denoted $P \geq 0$ if

$$x^T P x \geq 0 \qquad \text{for all } x \in \mathbb{R}^n$$

Definition 16.

A symmetric matrix $P \in \mathbb{S}^n$ is **Positive Definite**, denoted $P > 0$ if

$$x^T P x > 0 \qquad \text{for all } x \neq 0$$

- P is **Negative Semidefinite** if $-P \geq 0$
- P is **Negative Definite** if $-P > 0$
- A matrix which is neither Positive nor Negative Semidefinite is **Indefinite**

The set of positive or negative matrices is a *convex cone*.

Positive Matrices

Lemma 17.

$P \in \mathbb{S}^n$ *is positive definite if and only if all its eigenvalues are positive.*

Proof.

($\Rightarrow$) To show sufficiency, use the spectral decomposition. For $x \neq 0$,

$$x^T P x = x^T U^T \Lambda U x = v^T \Lambda v = \sum_i \lambda_i v_i^2 > 0$$

($\Leftarrow$) To show necessity, Suppose λ is an eigenvalue and $Px = \lambda x$ for $x \neq 0$, then

$$\lambda x^T x = x^T P x > 0$$

Hence

$$\lambda = \frac{x^T P x}{x^T x} > 0$$

□

Easy Proofs:

- A Positive Definite matrix is invertible.
- The inverse of a positive definite matrix is positive definite.

Positive Matrices

Easy Proofs:

- If $P > 0$, then $TPT^T \geq 0$ for any T. If T is invertible, then $TPT^T > 0$.

Lemma 18.

For any $P > 0$, there exists a positive square root, $P^{\frac{1}{2}} > 0$ such that $P = P^{\frac{1}{2}} P^{\frac{1}{2}}$.

Proof.

By the spectral decomposition $A = U \Lambda U^T$, where $\Lambda = \text{diag}(\lambda_1, \cdots, \lambda_n)$.

- Let $\Lambda^{\frac{1}{2}} = \text{diag}(\sqrt{\lambda_1}, \cdots, \sqrt{\lambda_n})$.
- Then $\Lambda^{\frac{1}{2}} > 0$, $\Lambda = \Lambda^{\frac{1}{2}} \Lambda^{\frac{1}{2}}$, and

$$\begin{aligned} A &= U \Lambda U^T \\ &= U \Lambda^{\frac{1}{2}} \Lambda^{\frac{1}{2}} U^T \\ &= U \Lambda^{\frac{1}{2}} U^T U \Lambda^{\frac{1}{2}} U^T \end{aligned}$$

- Thus $P = P^{\frac{1}{2}} P^{\frac{1}{2}}$, where $P^{\frac{1}{2}} = U \Lambda^{\frac{1}{2}} U^T > 0$

□

The properties of positive matrices extend to positive operators.

What is an inequality? What does ≥ 0 mean?

- An inequality implies a **partial ordering**:
  - $x \geq y$ if $x - y \geq 0$
- Any convex cone, C defines a partial ordering:
  - $x - y \geq 0$ if $x - y \in C$
- The ordering is only partial because $x \not\leq 0$ does not imply $x \geq 0$
  - $-x \notin C$ does not imply $x \in C$.
  - x may be indefinite.

Positivity will not change with coordinate transformations.

- $P > 0$ if and only if $T^*PT > 0$
- What about $T^{-1}PT$?

Theorem 19 (Schur Complement).

For any $S \in \mathbb{S}^n$, $Q \in \mathbb{S}^m$ and $R \in \mathbb{R}^{n\times m}$, the following are equivalent.

1. $\begin{bmatrix} M & R \\ R^T & Q \end{bmatrix} > 0$
2. $Q > 0$ *and* $M - RQ^{-1}R^T > 0$

Proof.

First, we show that 2) implies 1). Suppose that $Q > 0$ and $M - RQ^{-1}R^T > 0$. Then

$$\begin{bmatrix} M - RQ^{-1}R^T & 0 \\ 0 & Q \end{bmatrix} > 0$$

Thus

$$\begin{bmatrix} M & R \\ R^T & Q \end{bmatrix} = \begin{bmatrix} I & RQ^{-1} \\ 0 & I \end{bmatrix} \begin{bmatrix} M - RQ^{-1}R^T & 0 \\ 0 & Q \end{bmatrix} \begin{bmatrix} I & RQ^{-1} \\ 0 & I \end{bmatrix}^T > 0$$

Schur Complement

Proof.

We first show that 1) implies 2). Suppose that

$$\begin{bmatrix} M & R \\ R^T & Q \end{bmatrix} > 0$$

Then for any $x \in \mathbb{R}^m$, $x \neq 0$,

$$\begin{bmatrix} 0 \\ x \end{bmatrix}^T \begin{bmatrix} M & R \\ R^T & Q \end{bmatrix} \begin{bmatrix} 0 \\ x \end{bmatrix} = x^T Q x > 0.$$

Thus $Q > 0$ and hence Q^{-1} exists. Now, since $\begin{bmatrix} M & R \\ R^T & Q \end{bmatrix} > 0$,

$$\begin{bmatrix} I & -RQ^{-1} \\ 0 & I \end{bmatrix} \begin{bmatrix} M & R \\ R^T & Q \end{bmatrix} \begin{bmatrix} I & -RQ^{-1} \\ 0 & I \end{bmatrix}^T = \begin{bmatrix} M - RQ^{-1}R^T & 0 \\ 0 & Q \end{bmatrix} > 0$$

Therefore $M - RQ^{-1}R^T > 0$ □

Definition 20.

The **Orthogonal Complement** (Perp) of a subspace, $S \subset X$, is denoted

$$S^{\perp} := \left\{ x \in \mathbb{R}^n \;:\; \langle x, y \rangle = x^T y = 0 \qquad \text{for all } y \in S \right\}$$

Properties

- $\dim(S^{\perp}) = n - \dim(S)$
- For any $x \in \mathbb{R}^n$,

$$x = x_S + x_{S^{\perp}} \quad \text{for } x_S \in S \text{ and } x_{S^{\perp}} \in S^{\perp}$$

  - x_S and $x_{S^{\perp}}$ are unique.

Definition 21.

The orthogonal projection operator P_S is defined by

$$x_S = Px$$

if $x_S \in S$ and $x - x_S \in S^{\perp}$.

Projection and Perp space

P is a projection if and only if

$$P^2 = P$$

A projection P is orthogonal if and only if $P = P^*$

- Need Px and $(I - P)y$ orthogonal for all $x, y \in \mathbb{R}^n$.

Generalizes to any Hilbert space

Theorem 22.

For any $M \in \mathbb{R}^{n \times m}$, $[\mathit{Im}(M)]^\perp = \mathit{Ker}\left[M^T\right]$.

Proof.

We need to show $[\text{Im}(M)]^\perp \subset \text{Ker}\left[M^T\right]$ and $\text{Ker}\left[M^T\right] \subset [\text{Im}(M)]^\perp$.

- Suppose $x \in [\text{Im}(M)]^\perp$. If $x^T y = 0$ for any $y \in \text{Im}[M]$, then $x^T M z = 0$ for all z.
- Thus $z^T M^T x$ for all z. Let $z = M^T x$.
- Then $x^T M M^T x = \|M^T x\|^2 = 0$.
- Thus $x \in \text{Ker}\left[M^T\right]$, which implies $[\text{Im}(M)]^\perp \subset \text{Ker}\left[M^T\right]$.

□

Controllability

Proof.

We need to show Ker $\left[M^T\right] \subset [\text{Im}(M)]^{\perp}$.

- Suppose $x \in$ Ker $\left[M^T\right]$. Then

$$y^T M^T x = x^T M y = x^T z = 0$$

for any $z \in \text{Im}(M)$. Thus $x \in [\text{Im}(M)]^{\perp}$.

- This proves that $[\text{Im}(M)]^{\perp} =$ Ker $\left[M^T\right]$.

□