DEFINITIONS: Let $Q = \bar{x}^T A x$

1. Q (or A) is ***positive definite*** iff : $\langle x, Ax \rangle > 0$ for all $x \neq 0$.
2. Q (or A) is ***positive semidefinite*** if : $\langle x, Ax \rangle \geq 0$ for all $x \neq 0$.
3. Q (or A) is ***negative definite*** iff: $\langle x, Ax \rangle < 0$ for all $x \neq 0$.
4. Q (or A) is ***negative semidefinite*** if: $\langle x, Ax \rangle \leq 0$ for all $x \neq 0$.
5. Q (or A) is ***indefinite*** if: $\langle x, Ax \rangle > 0$ for some $x \neq 0$, and
   $\langle x, Ax \rangle < 0$ for other $x \neq 0$.

Tests for definiteness of matrix A in terms of its eigenvalues λ_i:

Matrix A is . . .	If the ***real parts*** of eigenvalues λ_i of A are: . . . .
1. Positive definite	All > 0
2. Positive semidefinite	All ≥ 0
3. Negative definite	All < 0
4. Negative semidefinite	All ≤ 0
5. Indefinite	Some $\mathrm{Re}(\lambda_i) > 0$, some $\mathrm{Re}(\lambda_i) < 0$.

There are also other tests involving *leading principal minors*.

We need to consider functions of matrices before we can solve the state equations in time domain.

Applying a function $f(A)$ to a matrix A is NOT the same thing as applying the function to the matrix entries element-by-element.

First, define matrix powers:

$$AA = A^2, \ldots \text{ etc.}$$

$$A^0 = I$$

$$A^m A^n = A^{m+n}$$

$$(A^m)^n = A^{mn}$$

$$(A^{-1})^n = A^{-n}$$

Matrix Polynomials:

polynomial form

Matrix Form	Scalar Form
$P(A) = c_m A^m + \cdots + c_1 A + c_0 I$	$P(x) = c_m x^m + \cdots + c_1 x + c_0$
$P(A) = c(A - Ia_1)\cdots(A - Ia_m)$	$P(x) = c(x - a_1)\cdots(x - a_m)$

factored form

Convergence of Polynomial Series:

Theorem: Let A be an n x n matrix whose eigenvalues are λ_i . If the infinite series

$$\sigma(x) = a_0 + a_1 x + a_2 x^2 + \cdots + a_k x^k + \cdots = \sum_{i=1}^{\infty} a_k x^k$$

converges for all $x = \lambda_i$, then . . .

. . . the series

$$\sigma(A) = a_0 I + a_1 A + a_2 A^2 + \cdots + a_k A^k + \cdots = \sum_{i=1}^{\infty} a_k A^k$$

converges. This will be important when we want the Taylor series expansions of a function of a matrix.

<u>Theorem</u>: If $f(z)$ is any function (not necessarily a polynomial) whose derivative exists for all z within a circle of the complex plane in which all eigenvalues of matrix A lie, then $f(A)$ can be written as a convergent power series.

<u>Example</u>: Find $\frac{d}{dt}\left(e^{At}\right)$

$$e^{At} = I + At + \frac{A^2t^2}{2!} + \cdots$$

$$\frac{de^{At}}{dt} = A + \frac{2A^2t}{2!} + \frac{3A^3t^2}{3!} + \cdots$$

$$= A\left[I + At + \frac{A^2t^2}{2!} + \cdots\right]$$

$$= Ae^{At} (= e^{At}A)$$

Also note that:

$$\sin(A) = A - \frac{A^3}{3!} + \frac{A^5}{5!} - \cdots$$

$$\cos(A) = I - \frac{A^2}{2!} + \frac{A^4}{4!} - \cdots$$

. . . etc., same as for expansions of scalar functions.

A much more useful theorem:

<u>Theorem</u>: Let $g(\lambda)$ be a polynomial of degree n-1 and $f(\lambda)$ be ANY function of λ. If $f(\lambda) = g(\lambda)$ for all eigenvalues of *A* ("on the spectrum of A"), then $f(A) = g(A)$ (for *A* itself.)

Implication: We can define the matrix-version of a non-polynomial scalar function using a matrix polynomial, if the two functions agree on the spectrum of the matrix!

<u>Example</u>: Let $A = \begin{bmatrix} 1 & 2 \\ 0 & 2 \end{bmatrix}$

"spectrum of A"={eigenvalues(A)}= $\sigma(A)$ ={1, 2}

Let $g(\lambda)$ be our n-1 order polynomial: $g(\lambda) = \alpha_0 + \alpha_1 \lambda$

Now suppose we are asked to find $f(A) = A^5$

$$f(\lambda) = \lambda^5$$

So we set $f(\lambda) = g(\lambda)$ for $\lambda = \{1,2\}$

$$\text{find } \alpha_0, \alpha_1 \quad \begin{cases} 1^5 = \alpha_0 + \alpha_1 \cdot 1 \\ 2^5 = 32 = \alpha_0 + \alpha_1 \cdot 2 \end{cases}$$

NOTE: If we had repeated eigenvalues, these equations would not be independent. We could instead use the equation AND its derivatives.

Solving, $\alpha_0 = -30, \quad \alpha_1 = 31$

Using this result:

$$A^5 = -30I + 31A = \begin{bmatrix} 1 & 62 \\ 0 & 32 \end{bmatrix}$$

Alternative way of calculating A

<u>Example</u>: Let $A = \begin{bmatrix} -3 & 1 \\ 0 & -2 \end{bmatrix}$ Find a *closed-form* solution for $\sin(A)$. (Can't use Taylor series) $\lambda_1 = -3, \lambda_2 = -2$

(goes on forever)

See previous theorem

This is similar to an earlier example. Because n=2, any analytic function of *A* can be written as a *first* order matrix polynomial, so

n-1 $\xrightarrow{\text{order}}$ $\sin(A) = \alpha_0 I + \alpha_1 A$

Evaluate this expression on the spectrum of *A*:

$$\left.\begin{aligned} \sin(-3) &= \alpha_0 + \alpha_1(-3) \\ \sin(-2) &= \alpha_0 + \alpha_1(-2) \end{aligned}\right\} \quad \begin{aligned} \alpha_1 &= -0.768 \\ \alpha_0 &= -2.45 \end{aligned}$$

Solving,

$$\sin(A) = \begin{bmatrix} \sin(-3) & \sin(-2) - \sin(-3) \\ 0 & \sin(-2) \end{bmatrix} = \alpha_0 I + \alpha_1 A$$

If *A* had repeated eigenvalues, the two equations

$$\sin(-3) = \alpha_0 + \alpha_1(-3)$$
$$\sin(-2) = \alpha_0 + \alpha_1(-2)$$

would be linearly dependent and have no unique solution. Then we could use one of them and use a *derivative* for the other:

$$\frac{d}{d\lambda}\left[\sin(\lambda) = \alpha_0 + \alpha_1\lambda\right]$$

so

$$\cos(\lambda) = \alpha_1$$

This would be the second independent equation.

<u>Cayley-Hamilton Theorem:</u> Let a system have characteristic polynomial

$$\left|A - \lambda I\right| = \phi(\lambda)$$

Then

$$\phi(A) = 0$$

That is, every matrix satisfies its own characteristic polynomial.

This theorem, together with the previous one, imply that we never need to consider polynomials of a matrix of order higher than *n-1* (!!)

<u>Example</u>: (Reduction of matrix polynomials to degree n-1 or less). Let

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

and find $P(A) = A^4 + 3A^3 + 2A^2 + A + I.$ } characteristic equation

$$\Delta(\lambda) = \lambda^2 - 5\lambda + 5 = 0$$

$$A^2 = -5I + 5A$$

So from Cayley-Hamilton,

$$A^4 = A^2 A^2 = (5A - 5I)^2 = 25A^2 - 50A + 25I = 25(5A - 5I) - 50A + 25I$$

$$A^3 = A^2 A = (5A - 5I)A = 5A^2 - 5A = 5(5A - 5I) - 5A$$

$$A^2 = 5A - 5I$$

Now *P(A)* will contain no powers of A higher than 1.

(=n-1)

Some examples of what these theorems allow us to do:

<u>Example</u>: Suppose the characteristic polynomial of a system is

$$\phi(\lambda) = \lambda^n + c_{n-1}\lambda^{n-1} + \cdots + c_1\lambda + c_0 = 0$$

so

$$\phi(A) = A^n + c_{n-1}A^{n-1} + \cdots + c_1 A + c_0 I = 0$$

Noting that c_0 is equal to the product of all the eigenvalues, we know it is nonzero iff matrix *A* is non-singular (no zero eigenvalues), or *A* is invertible. Multiply the above equation through by A^{-1} to get:

$$A^{n-1} + c_{n-1}A^{n-2} + \cdots + c_1 I + c_0 A^{-1} = 0$$

Solving $$A^{-1} = -\frac{1}{c_0}\left[A^{n-1} + c_{n-1}A^{n-2} + \cdots + c_1 I\right]$$ Easy way for computer to find inverse

<u>Definition</u>: The minimal polynomial of a square matrix A is the lowest degree monic polynomial $\phi_m(\lambda)$ which satisfies

$$\phi_m(A) = 0$$

Being minimal affects only powers of repeated terms in characteristic polynomials, for example, if

$$\phi(\lambda) = (\lambda - \lambda_1)^{m_1}(\lambda - \lambda_2)^{m_2} \cdots (\lambda - \lambda_p)^{m_p},$$

$$\phi_m(\lambda) = (\lambda - \lambda_1)^{\eta_1}(\lambda - \lambda_2)^{\eta_2} \cdots (\lambda - \lambda_p)^{\eta_p}$$

$$\eta_i \le m_i$$

where

How many ones in super diagonal of Jordan Form

Note that η_i is not necessarily 1, but is rather the *index* of the eigenvalue λ_i.

Another important example of this technique will be in the computation of the matrix exponential:

$$e^{At}$$

We will see in the next chapter how important this matrix will be in the solution of the state variable equations for a system.

Solutions to State Equations

We know how to solve scalar linear differential equations, but what about the state-space equations:

$$\dot{x} = Ax + Bu, \qquad x(t_0) = x_0$$
$$y = Cx + Du$$

Actually, we need only to consider $\dot{x} = Ax + Bu$ because finding y will then be a simple matter of matrix multiplication.

Recall the technique of *integrating factor* in the solution of linear differential equations:

$$\dot{x} - Ax = Bu$$

Multiplying this equation by e^{-At} will result in the left-hand side becoming a "perfect" differential:

$$e^{-At}\left[\dot{x} - Ax = Bu\right]$$

$$e^{-At}\dot{x} - e^{-At}Ax = e^{-At}Bu$$

$$\frac{d}{dt}\left[e^{-At}x(t)\right] = e^{-At}Bu(t)$$

Now multiply both sides by dt and integrate over a dummy variable τ from t_0 to t .

$$e^{-At}x(t) - e^{-At_0}x(t_0) = \int_{t_0}^{t} e^{-A\tau}Bu(\tau)d\tau$$

$$e^{-At}x(t) - e^{-At_0}x(t_0) = \int_{t_0}^{t} e^{-A\tau}Bu(\tau)d\tau$$

Move initial condition term to RHS and multiply through by e^{At}

$$x(t) = e^{At}e^{-At_0}x(t_0) + e^{At}\int_{t_0}^{t} e^{-A\tau}Bu(\tau)d\tau$$

$$= e^{A(t-t_0)}x(t_0) + \int_{t_0}^{t} e^{A(t-\tau)}Bu(\tau)d\tau$$

Note that if matrix *B* were a function of time, this would become simply

$$\boxed{x(t) = e^{A(t-t_0)}x(t_0) + \int_{t_0}^{t} e^{A(t-\tau)}B(\tau)u(\tau)d\tau}$$

(but if *A* were a function of time, we run into bigger problems.)

If we wanted to compute $y(t)$, we would simply get:

$$y(t) = Ce^{A(t-t_0)}x(t_0) + C\int_{t_0}^{t} e^{A(t-\tau)}Bu(\tau)d\tau + Du(t)$$

Again, *C* and *D* could be functions of time without complicating matters too much. If *A* is time-varying, we must be more careful choosing a proper integrating factor. The matrix exponential will no longer work.

Again, the importance of the matrix exponential e^{At} arises. We'll summarize the several ways to compute it shortly.

System modes and modal decompositions:

This is a very powerful representation of a system's solutions, used widely in large-scale systems and infinite-dimensional systems, which are often represented by partial differential equations rather than ordinary differential equations. It underscores the importance of a basis of the state-space.

Let the set $\{\xi_i\}$ be the set of *n* linearly independent eigenvectors, including, if necessary, generalized eigenvectors, corresponding to eigenvalues λ_i of the **constant** matrix *A*. Because this set forms a basis of the state-space, we can write

$$x(t) = \sum_{i=1}^{n} q_i(t)\xi_i$$

$q_i(t)$ denotes the scalar coefficients

ξ_i denotes the the eigenvectors

$$x(t) = \sum_{i=1}^{n} q_i(t)\xi_i$$

For some time-varying coefficients

$$q_i(t), \qquad i = 1, \ldots, n$$

We can easily do the same for the term $B(t)u(t)$

$$B(t)u(t) = \sum_{i=1}^{n} \beta_i(t)\xi_i$$

$\beta_i(t)$ denotes the scalar coefficients

ξ_i eigenvectors

Substituting these expansions into the state-equations

$$\dot{x} = Ax + Bu$$

Gives . . . . .

$$\sum_{i=1}^{n} \dot{q}_i(t)\xi_i = \sum_{i=1}^{n} q_i(t)A\xi_i + \sum_{i=1}^{n} \beta_i(t)\xi_i$$

We have implicitly assumed in this step that we have *n* linearly independent eigenvectors. If this is not the case, relatively minor complications arise.

Re-arranging, $$\sum_{i=1}^{n} \left(\dot{q}_i(t) - q_i(t)A - \beta_i(t)\right)\xi_i = 0$$

These coefficients must **all** be zero, so

$$\dot{q}_i(t) = q_i(t)\lambda_i + \beta_i(t) \quad \text{for} \quad i = 1, \ldots, n$$

Recall $\boxed{A\xi_i = \lambda_i \xi_i}$ } Eigenvalue/Eigenvector Problem

$$\dot{q}_i(t) = q_i(t)\lambda_i + \beta_i(t) \quad \text{for} \quad i = 1, \ldots, n$$

This is a set of *n* **de-coupled** equations (if we had used any generalized eigenvectors, some would still be coupled, but only to one other equation). } Jordan Form

The terms $q_i(t)\xi_i$ are called ***system modes***, and are equivalent to the "new" state variables $\bar{x}(t)$ that we obtained in the past example where we "diagonalized" the system using the modal matrix. Recall that if *M* is the modal matrix, we can define new variables.

$$x = Mq$$

such that

$$\dot{q} = M^{-1}AMq + M^{-1}Bu$$
$$y = CMq + Du$$

$$\dot{q} = M^{-1}AMq + M^{-1}Bu$$
$$y = CMq + Du$$

where $M^{-1}AM = J$ is the *Jordan form* of the *A*-matrix (diagonal if there are n eigenvectors).

Because these equations are decoupled, the solutions to the state equations are particularly simple. We can find the solutions $q(t)$ and then change them back to the original variables $x(t)$ by un-doing the transformation afterward. That is,

$$J = M^{-1}AM$$

Initial Conditions

$$q(t) = e^{J(t-t_0)}q(t_0) + \int_{t_0}^{t} e^{J(t-\tau)}M^{-1}B(\tau)u(\tau)d\tau, \boxed{q(t_0) = M^{-1}x(t_0)}$$

after which

$$x(t) = Mq(t)$$

In ***diagonal*** form, the computation of e^{Jt} is particularly easy:

$$e^{\begin{bmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{bmatrix} t} = e^{\begin{bmatrix} \lambda_1 t & & 0 \\ & \ddots & \\ 0 & & \lambda_n t \end{bmatrix}}$$

and

$$e^{At} = Me^{Jt}M^{-1}$$

Whenever two matrices *A* and *J* are similar, we can compute our function of *J* and perform the reverse-similarity transform afterward. That is,

if $J = M^{-1}AM$,

then $f(A) = Mf(J)M^{-1}$ and $f(J) = M^{-1}f(A)M$

Note that $\hat{A}$ and A are similar if

$\hat{A} = M^{-1}AM$ for some orthonormal M

Often in large-scale or infinite-dimensional systems, some modes are negligible and are discarded after modal expansion, thus reducing the size of the system. For example, when a beam vibrates, we have an infinite number of terms in a series expansion of its displacement function, but only the first few (2 - 5) may dominate.