Cayley-Hamilton Theorem

Every square matrix A satisfies its own characteristic equation:
$\Delta(A) = 0$
where the characteristic equation (aka characteristic polynomial) is given by:
$\Delta(\lambda) = |\lambda I - A| = \lambda^n + c_{n-1}\lambda^{n-1} + c_{n-2}\lambda^{n-2} + ...c_1\lambda + c_0 = 0.$

Proof (For case when A is similar to a diagonal matrix, i.e. for $A \in \Re^{n\times n}$ with $A = P\Lambda P^{-1}$ where Λ is a diagonal matrix with elements on the diagonal $\lambda_1, \lambda_2, ...\lambda_n$.)

Substituting A in the characteristic polynomial, we have

$$\Delta(A) = A^n + c_{n-1}A^{n-1} + c_{n-2}A^{n-2} + ...c_1 A + c_0 I \tag{1}$$

Noting that $A^k = P\Lambda^k P^{-1}$, then

$$\Delta(A) = P[\Lambda^n + c_{n-1}\Lambda^{n-1} + c_{n-2}\Lambda^{n-2} + ...c_1\Lambda + c_0 I]P^{-1}. \tag{2}$$

Since Λ is diagonal, the typical i, i term is given by

$$\Delta(\lambda_i) = |\lambda_i I - A| = \lambda_i^n + c_{n-1}\lambda_i^{n-1} + c_{n-2}\lambda_i^{n-2} + ...c_1\lambda_i + c_0 = 0.$$

Where the sum is zero because λ_i is a root of the characteristic polynomial. Thus $\Delta(A) = P[0]P^{-1} = [0] \quad \square$.

Matrix Exponential

Recall series form for $e^{At} = I + At + A^2\frac{t^2}{2!} + A^3\frac{t^3}{3!} + ...$ But from Cayley-Hamilton, we know that since $\Delta(A) = 0$ then $-A^n = c_{n-1}A^{n-1} + c_{n-2}A^{n-2} + ...c_1 A + c_0 I$. And then all higher powers than A^n can be expressed in terms of a linear sum of $I, A, A^2, ..., A^{n-1}$.

Then
$e^{At} = \alpha_0(t)I + \alpha_1(t)A + ...\alpha_{n-1}(t)A^{n-1} = R(A)$
where for a given t, $R(A)$ is a polynomial of degree $n-1$,
and α_i are found by solving $e^{\lambda_i t} = \alpha_0(t) + \alpha_1(t)\lambda_i + ...\alpha_n(t)\lambda_i^n$.

Example of using C-H for matrix exponential. Given

$$A = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix} \tag{3}$$

The matrix exponential can be calculated easily using Laplace Transform:

$$e^{At} = \mathcal{L}^{-1}[sI - A]^{-1} = \begin{bmatrix} e^t & 0 \\ 0 & e^{2t} \end{bmatrix}. \tag{4}$$

The matrix exponential can alternatively be calculated by Cayley-Hamilton: $e^{At} = \alpha_0(t)I + \alpha_1(t)A$. The functions $\alpha_i(t)$ are found using $\lambda_1 = 1, \lambda_2 = 2$ by solving

$$\begin{aligned} e^{\lambda_1 t} = e^t &= \alpha_0(t) + 1\alpha_1(t) && (5)\\ e^{\lambda_2 t} = e^{2t} &= \alpha_0(t) + 2\alpha_1(t) && (6)\\ & && (7) \end{aligned}$$

Thus $\alpha_1(t) = e^{2t} - e^t$ and $\alpha_0(t) = 2e^t - e^{2t}$. Finally,

$$e^{At} = (2e^t - e^{2t})\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + (e^{2t} - e^t)\begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix} = \begin{bmatrix} e^t & 0 \\ 0 & e^{2t} \end{bmatrix} \tag{8}$$

Controllability

Assume that $\dot{\mathbf{x}} = A\mathbf{x} + B\mathbf{u}$ is completely controllable. Recall that

$$\mathbf{x}(t) = e^{A(t-t_o)}\mathbf{x}(t_o) + \int_{t_o}^{t} e^{A(t-\tau)}Bu(\tau)d\tau. \tag{9}$$

Since by assumption the system is controllable, we can choose a final time t_1 such that $\mathbf{x}(t_1) = \mathbf{0}$ with initial condition $\mathbf{x}(0) = \mathbf{x_0}$ with $t_o = 0$. So by eqn(9) we have

$$-\mathbf{x_o} = \int_{0}^{t_1} e^{-A\tau}Bu(\tau)d\tau. \tag{10}$$

By Cayley-Hamilton, we can express $e^{-A\tau}$ as a polynomial in A:

$$e^{-A\tau} = \alpha_0(\tau)I + \alpha_1(\tau)A + \alpha_2(\tau)A^2 + ... + \alpha_{n-1}(\tau)A^{n-1} = \sum_{j=0}^{n-1} A^j\alpha_j(\tau). \tag{11}$$

If we substitute eqn(11) into eqn(10) we obtain

$$-\mathbf{x_o} = \sum_{j=0}^{n-1} A^j B \int_{0}^{t_1} \alpha_j(\tau)u(\tau)d\tau. \tag{12}$$

Note that $\int_0^{t_1} \alpha_j(\tau)u(\tau)d\tau$ is a constant. Define $v_j = \int_0^{t_1} \alpha_j(\tau)u(\tau)d\tau$. Then eqn. (12) can be expressed as a matrix multiply:

$$-\mathbf{x_o} = \left[B|AB|A^2B|...|A^{n-1}B\right] \begin{bmatrix} v_0 \\ v_1 \\ ... \\ v_{n-1} \end{bmatrix}. \tag{13}$$

Define the *controllability matrix* $\mathbb{C} = \left[B|AB|A^2B|...|A^{n-1}B\right]$. Note that if state space is of dimension n, then eqn(13) will only be satisfiable for all $\mathbf{x_o}$ if rank $(\mathbb{C}) = n$. Thus the necessary condition for controllability is shown. $\square$