1 More on the Cayley-Hamilton Theorem

1.1 How to evaluate polynomial functions of a matrix?

**Problem:**
Given \( p(s) \) of order \( m \geq n \) evaluate \( p(A) \) for some matrix \( A \) of order \( n \).

**First answer:**
Compute the characteristic polynomial \( d_A(s) \) of \( A \) with order \( n \). Then use the Euclidian algorithm for polynomial division to write

\[
p(s) = q(s)d_A(s) + r(s)
\]

where \( r(s) \) has degree at most \( n - 1 \). From the Cayley-Hamilton Theorem \( d_A(A) = 0 \), so that

\[
p(A) = q(A)d_A(A) + r(A) = r(A).
\]

**Example:**
Compute \( A^5 + A^3 \) for \( A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \).

For this problem \( p(s) = s^5 + s^3 \) and \( d_A(s) = (s - 1)^2 = s^2 - 2s + 1 \). Therefore

\[
\overbrace{s^5 + s^3}^{p(s)} = \overbrace{(s^3 + 2s^2 + 4s + 6)}^{q(s)} \overbrace{(s^2 - 2s + 1)}^{d_A(s)} + \overbrace{(8s - 6)}^{r(s)},
\]

and

\[
A^5 + A^3 = p(A) = r(A) = 8A - 6I = 8 \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} - 6 \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 2 & 8 \\ 0 & 2 \end{bmatrix}.
\]
Second answer:
Since there exists
\[ p(s) = q(s)d_A(s) + r(s) \]
where \( r(s) \) is of degree at most \( n - 1 \) and if \( \lambda_i, i = 1, \ldots, n \) are the eigenvalues of \( A \) then
\[ p(\lambda_i) = q(\lambda_i)d_A(\lambda_i) + r(\lambda_i) = r(\lambda_i), \quad \forall i = 1, \ldots, n. \]
The above gives us \( n \) equations on \( n \) unknowns (the coefficients of \( r! \)).

Example:
Compute \( A^{1000} \) for \( A = \begin{bmatrix} 1 & 1 \\ 0 & 2 \end{bmatrix} \). For this problem \( p(s) = s^{1000} \) and \( \lambda_1 = 1, \lambda_2 = 2 \). Therefore for \( r(s) = r_1s + r_2 \)
\[ r_1 + r_2 = r_1 \lambda_1 + r_2 = r(\lambda_1) = p(\lambda_1) = \lambda_1^{1000} = 1, \]
\[ 2r_1 + r_2 = r_1 \lambda_2 + r_2 = r(\lambda_2) = p(\lambda_2) = \lambda_2^{1000} = 2^{1000}. \]
Solving the above equations we have
\[ r_1 = 2^{1000} - 1, \]
\[ r_2 = 1 - r_1 = 2 - 2^{1000}, \]
and
\[ A^{1000} = r(A) = (2^{1000} - 1) \begin{bmatrix} 1 & 1 \\ 0 & 2 \end{bmatrix} + (2 - 2^{1000}) \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 2^{1000} - 1 \\ 0 & 2^{1000} - 1 \end{bmatrix}. \]
1.2 Interpolation

Let \( d_A(s) = \prod_{i=1}^{m} (s - \lambda_i)^{k_i}, \) \( \sum_{i=1}^{m} k_i = n \) and \( f(s) \) be a function with at least \( r - 1 \) derivatives, where \( r = \max_i k_i. \) The polynomial
\[
r(s) = r_1 s^{n-1} + r_2 s^{n-2} + \cdots + r_n
\]
is said to interpolate \( f \) and its derivatives at the roots of \( d_A \) if
\[
f^{(j-1)}(\lambda_i) = r^{(j-1)}(\lambda_i), \quad \forall i = 1, \ldots, m, \quad j = 1, \ldots, k_i.
\]

**Proposition:** When \( f(s) \) is a polynomial of degree \( m \) and \( r(s) \) is a polynomial of degree \( q(s) \) are polynomials such that
\[
f(s) = q(s)d_A(s) + r(s),
\]
then \( r \) interpolates \( f \) and its derivatives at the roots of \( d_A. \)

**Proof:** For all \( \lambda_i, i = 1, \ldots, m, j = 1, \)
\[
f(\lambda_i) = q(\lambda_i)d_A(\lambda_i) + r(\lambda_i) = r(\lambda_i).
\]
Note that
\[
f'(s) = q'(s)d_A(s) + q(s)d'_A(s) + r'(s),
\]
and for all \( i \) such that \( k_i > 1 \) we have
\[
d'_A(\lambda_i) = 0
\]
so that
\[
f'(\lambda_i) = q'(\lambda_i)d_A(\lambda_i) + q(s)d'_A(\lambda_i) + r'(\lambda_i) = r'(\lambda_i).
\]
In general, for \( i \) such that \( k_i > 1 \) we have
\[
d_A^{(j-1)}(\lambda_i) = 0, \quad j = 1, \ldots, k_i,
\]
which implies
\[
f^{(j-1)}(\lambda_i) = r^{(j-1)}(\lambda_i), \quad j = 1, \ldots, k_i.
\]
1.3 How to evaluate non-polynomial functions of a matrix?

**Problem:**
Given $f(s)$ with at least $r-1$ derivatives and a matrix $A$ of order $n$, where $r$ is maximum multiplicity of the eigenvalues of $A$, evaluate $f(A)$.

**Answer:**
Compute the polynomial $r$ of degree $n-1$ that interpolates $f$ at the roots of $d_A$. Then $f(A) = r(A)$.

**Example:**
Compute $e^A$ for $A = \begin{bmatrix} 1 & 1 \\ 0 & 2 \end{bmatrix}$. For this problem $\lambda_1 = 1$, $\lambda_2 = 2$. Therefore, for $r(s) = r_1 s + r_2$

\[
\begin{align*}
    r_1 + r_2 &= r_1 \lambda_1 + r_2 = r(\lambda_1) = f(\lambda_1) = e^{\lambda_1} = e, \\
    2r_1 + r_2 &= r_1 \lambda_2 + r_2 = r(\lambda_2) = f(\lambda_2) = e^{\lambda_2} = e^2
\end{align*}
\]

Solving the above equations we have

\[
\begin{align*}
    r_1 &= e^2 - e, \\
    r_2 &= e - r_1 = 2e - e^2
\end{align*}
\]

and

\[
e^A = r(A) = (e^2 - e) \begin{bmatrix} 1 & 1 \\ 0 & 2 \end{bmatrix} + (2e - e^2) \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} e & e^2 - e \\ 0 & e^2 \end{bmatrix}.
\]
2 More on Controllability and Observability

2.1 Non-controllable realizations

Assume \((A, B)\) is not controllable and that \(\text{rank} C(A, B) = r < n\)

**Proposition:** There exist a nonsingular matrix \(T\) such that

\[
\bar{A} = T^{-1}AT = \begin{bmatrix} A_c & A_c \bar{c} \\ 0 & A_{\bar{c}} \end{bmatrix}, \quad \bar{B} = T^{-1}B = \begin{bmatrix} B_{\bar{c}} \end{bmatrix}, \quad \bar{C} = CT = \begin{bmatrix} C_{\bar{c}} & C_{\bar{c}} \end{bmatrix},
\]

where \(A_c \in \mathbb{R}^{r \times r}\) and \((A_c, B_c)\) is controllable.

**Proof:**

First note that

\[
C(\bar{A}, \bar{B}) = \begin{bmatrix} \bar{B} & \bar{A}\bar{B} & \cdots & \bar{A}^{r-1}\bar{B} & \cdots & \bar{A}^{n-1}\bar{B} \end{bmatrix},
\]

so that we can use the Cayley-Hamilton Theorem to show that \(C(\bar{A}, \bar{B})\) has rank \(r\) if and only if \(C(A_c, B_c)\) has rank \(r\). Furthermore

\[
C(\bar{A}, \bar{B}) = \begin{bmatrix} T^{-1}B & T^{-1}AB & \cdots & T^{-1}A^{n-1}B \end{bmatrix},
\]

therefore \(C(\bar{A}, \bar{B})\) must have rank \(r\), and so has \(C(A_c, B_c)\).

From the above

\[
C(A, B) = \begin{bmatrix} \bar{B} & \bar{A}\bar{B} & \cdots & \bar{A}^{r-1}\bar{B} & * \end{bmatrix},
\]

which implies that

\[
T_1 = \begin{bmatrix} \bar{B} & \bar{A}\bar{B} & \cdots & \bar{A}^{r-1}\bar{B} \end{bmatrix} C^\dagger(A_c, B_c),
\]

where the symbol \(X^\dagger\) denotes the pseudo-inverse of \(X\) (more on that later!).

As \(T_1\) has full rank, matrix \(T_2\) can be chosen to make \(T\) nonsingular.
Corollary: \( C(sI - A)^{-1}B = C_{\bar{c}}(sI - A_{\bar{c}})^{-1}B_{\bar{c}}. \)

Proof: Verify that

\[
\begin{bmatrix}
C_c & C_{\bar{c}} \\
\end{bmatrix}
\begin{bmatrix}
sI - A_c & -A_{c\bar{c}} \\
0 & sI - A_{\bar{c}} \\
\end{bmatrix}^{-1}
\begin{bmatrix}
B_c \\
\end{bmatrix}
= \begin{bmatrix}
C_c & C_{\bar{c}} \\
\end{bmatrix}
\begin{bmatrix}
(sI - A_c)^{-1} & * \\
0 & (sI - A_{\bar{c}})^{-1} \\
\end{bmatrix}^{-1}
\begin{bmatrix}
0 \\
\end{bmatrix},
\]

\[
= \begin{bmatrix}
C_c & C_{\bar{c}} \\
\end{bmatrix}
\begin{bmatrix}
(sI - A_c)^{-1}B_c \\
0 \\
\end{bmatrix},
\]

\[
= C_c(sI - A_c)^{-1}B_c.
\]

2.2 Non-observable realizations

Assume \((A, C)\) is not observable and that

\[
\text{rank } \mathcal{O}(A, c) = r < n
\]

Proposition: There exist a nonsingular matrix \(T\) such that

\[
\tilde{A} = T^{-1}AT = \begin{bmatrix}
A_o & 0 \\
A_{\bar{o}o} & A_{\bar{o}} \\
\end{bmatrix}, \quad \tilde{B} = T^{-1}B = \begin{bmatrix}
B_o \\
B_{\bar{o}} \\
\end{bmatrix}, \quad \tilde{C} = CT = \begin{bmatrix}
C_o & 0 \\
\end{bmatrix},
\]

where \(A_o \in \mathbb{R}^{r \times r}\) and \((A_o, C_o)\) is observable.
2.3 The Popov-Belevitch-Hautus Test

**Theorem:** The pair \((A, C)\) is observable if and only if there exists no \(x \neq 0\) such that
\[ Ax = \lambda x, \quad Cx = 0. \]  

**Proof:**

* Sufficiency: Assume there exists \(x \neq 0\) such that (1) holds. Then
\[ CAx = \lambda Cx = 0, \]
\[ CA^2x = \lambda CAx = 0, \]
\[ \vdots \]
\[ CA^{n-1}x = \lambda CA^{n-2}x = 0 \]
so that
\[ O(A, C)x = 0, \]
which implies that the pair \((A, C)\) is not observable.

* Necessity: Assume that \((A, C)\) is not observable. Then transform it into the equivalent non observable realization where
\[ \bar{A} = \begin{bmatrix} A_0 & 0 \\ A_\bar{o} & A_\bar{\bar{o}} \end{bmatrix}, \quad \bar{C} = \begin{bmatrix} C_0 & 0 \end{bmatrix}. \]
Chose \(x \neq 0\) such that
\[ A_\bar{o}x = \lambda x. \]

Then
\[ \begin{bmatrix} A_0 & 0 \\ A_\bar{o} & A_\bar{\bar{o}} \end{bmatrix} \begin{bmatrix} 0 \\ x \end{bmatrix} = \lambda \begin{bmatrix} 0 \\ x \end{bmatrix}, \quad \begin{bmatrix} C_0 & 0 \end{bmatrix} \begin{bmatrix} 0 \\ x \end{bmatrix} = 0. \]

**Theorem:** The pair \((A, B)\) is observable if and only if there exists no \(z \neq 0\) such that
\[ z^*A = \lambda z, \quad z^*B = 0. \]