A. Matrix algebra

A.1 The Caley-Hamilton theorem

This result states that every square matrix satisfies its own characteristic polynomial:

Theorem A.1.1 Let $A \in \mathbb{R}^{n \times n}$ have the characteristic polynomial

$$p(\lambda) = \det(\lambda I - A) = \lambda^n + \alpha_{n-1}\lambda^{n-1} + \cdots + \alpha_1 \lambda + \alpha_0.$$

Then $p(A) = 0$.

A.2 The matrix exponential

The matrix exponential is a matrix-valued function used for solving systems of differential equations. Analogous to how the solution to a scalar differential equation

$$\dot{x}(t) = ax(t)$$

is given in terms of the ordinary exponential function

$$x(t) = e^{at}x(0),$$

the solution to a system of linear equations

$$\dot{x}(t) = Ax(t) \tag{A.1}$$

with $x(t) \in \mathbb{R}^n$ is given by

$$x(t) = e^{At}x(0) \tag{A.2}$$

where e^{At} is the matrix exponential (of the matrix At).

Definition A.2.1 The matrix exponential of $M \in \mathbb{R}^{n\times n}$ is defined by the power series

$$e^M = I + M + \frac{1}{2!}M^2 + \frac{1}{3!}M^3 + \cdots$$

We can directly verify that the solution (A.2) satisfies (A.1):

$$\begin{aligned}\dot{x}(t) &= \frac{d}{dt}\left(I + At + \frac{1}{2!}A^2t^2 + \cdots\right)x(0) = \\ &= \left(A + A^2t + \frac{1}{2!}A^3t^2 + \cdots\right)x(0) = \\ &= A(I + At + \frac{1}{2!}A^2t^2 + \cdots)x(0) = Ae^{At}x(0) = Ax(t)\end{aligned}$$

For nilpotent matrices, the definition can also be used to evaluate the matrix exponential (since the series converges after a finite number of terms). In general, however, it is often more convenient to use the Laplace transform. To understand how this works, recall that

$$(I - M)^{-1} = I + M + M^2 + M^3 + \cdots$$

provided that the series converges (you can verify the identity by multiplying both sides of the equation with $(I - M)$). Thus,

$$(sI - A)^{-1} = \frac{1}{s}(I - \frac{A}{s})^{-1} = \frac{1}{s}I + \frac{1}{s^2}A + \frac{1}{s^3}A^2 + \cdots$$

which converes for large enough $|s|$. By the inverse Laplace transform

$$\mathcal{L}^{-1}\left((sI - A)^{-1}\right) = I + tA + \frac{t^2}{2!}A^2 + \cdots = e^{At}.$$

The next example demonstrates the two techniques for computing the matrix exponential.

■ **Example A.1** Consider

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

then since $A^k = 0$ for $k \geq 2$ (A is nilpotent) the power series allows us to conclude that

$$e^{At} = I + At = \begin{bmatrix} 1 & t \\ 0 & 1 \end{bmatrix}.$$

Using the inverse Laplace transform approach, we would first compute

$$(sI - A)^{-1} = \begin{bmatrix} s & -1 \\ 0 & s \end{bmatrix}^{-1} = \frac{1}{s^2}\begin{bmatrix} s & 1 \\ 0 & s \end{bmatrix} = \begin{bmatrix} 1/s & 1/s^2 \\ 0 & 1/s \end{bmatrix}.$$

Then we compute the matrix exponential by taking the inverse Laplace transform

$$e^{At} = \mathcal{L}^{-1}\left((sI - A)^{-1}\right) = \begin{bmatrix} \mathcal{L}^{-1}(1/s) & \mathcal{L}^{-1}(1/s^2) \\ 0 & \mathcal{L}^{-1}(1/s) \end{bmatrix} = \begin{bmatrix} 1 & t \\ 0 & 1 \end{bmatrix}$$

■

A.3 Matrix inversion lemmas

The following two matrix inversion lemmas are useful for our derivations. Although the expressions look complicated, they are trivial to prove: just multiply the original matrix expression with the formula for its inverse and verify that the product evaluates to the identity matrix.

Proposition A.3.1 Consider the matrix $A \in \mathbb{R}^{n\times n}$ partitioned as

$$A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}$$

where $A_{11} \in \mathbb{R}^{k\times k}$ and $A_{22} \in \mathbb{R}^{(n-k)\times(n-k)}$ are both invertible. Then

$$A^{-1} = \begin{pmatrix} (A_{11} - A_{12}A_{22}^{-1}A_{21})^{-1} & -A_{11}^{-1}A_{12}(A_{22} - A_{21}A_{11}^{-1}A_{12})^{-1} \\ -A_{22}^{-1}A_{21}(A_{11} - A_{12}A_{22}^{-1}A_{21})^{-1} & (A_{22} - A_{21}A_{11}^{-1}A_{12})^{-1} \end{pmatrix}$$

Proposition A.3.2 Let $X \in \mathbb{R}^{n\times n}$, $Y \in \mathbb{R}^{k\times k}$ and $Z \in \mathbb{R}^{n\times k}$ be real matrices of appropriate dimensions with X and Y invertible. Then

$$(X + ZYZ^T)^{-1} = X^{-1} - X^{-1}Z(Y^{-1} + Z^TX^{-1}Z)^{-1}Z^TX^{-1} \tag{A.3}$$

and

$$Y^{-1}Z^T(X^{-1} + ZY^{-1}Z^T)^{-1} = (Y + Z^TXZ)^{-1}Z^TX. \tag{A.4}$$

We note that Equation (A.3) is a special case of the Sherman-Morrison-Woodbury formula.