A.2 Differentiation of Matrices

Matrix-elements may be functions of one or more variables. These functions may be differentiable, in which case the derivative of the matrix may be defined.

• Definition. The derivative of a matrix A is a matrix whose elements are the derivative of the matrix-elements of A.

$$A' = \frac{dA(x)}{dx}, \quad \text{matrix-elements} \quad a'_{ij} = \frac{da_{ij}(x)}{dx}.$$

The derivative of a matrix will in general not commute with the matrix itself. Apart from this the usual equations from differential calculus are valid here too, as for example $d(A+B)/dx = dA/dx + dB/dx$. However,

$$\frac{d(A^2)}{dx} = \frac{dA}{dx} \cdot A + A \cdot \frac{dA}{dx}.$$

This is a consequence of the chain rule:

$$\frac{d}{dx} \sum_k a_{ik} a_{kj} = \sum_k a'_{ik} a_{kj} + \sum_k a_{ik} a'_{kj}.$$

In general:

$$\frac{dAB}{dx} = A'B + AB'.$$

The derivative of the inverse of a matrix A can be found as follows. One has $A^{-1}A = I$ and thus

$$0 = \frac{d}{dx}(A^{-1}A) = \frac{dA^{-1}}{dx} A + A^{-1} \frac{dA}{dx},$$

or

$$\frac{dA^{-1}}{dx} = -A^{-1} \frac{dA}{dx} A^{-1}.$$

A.3 Functions of Matrices

Functions based on addition and multiplication may be defined for matrices as well. Especially the function e^x is important. There are two equivalent definitions:

$$e^A = \sum_{k=0}^{\infty} \frac{1}{k!} A^k, \quad \text{with} \quad A^0 = I \tag{A.4}$$

$$e^A = \lim_{m \to \infty} (I + \frac{1}{m} A)^m. \tag{A.5}$$

The demonstration that the second definition is the same as the first is as usual, namely by application of the binomial expansion to the second equation.

Equation (A.5) allows the derivation of an interesting equation. Consider

$$Det(e^A) = \lim_{m \to \infty} \left[Det(I + \frac{1}{m} A) \right]^m . \tag{A.6}$$

The determinant inside the square brackets may be computed, ignoring terms of order $1/m^2$ or smaller. Consider for example a 2×2 matrix. Then:

$$I + \frac{1}{m} A = \begin{pmatrix} 1 + \frac{a_{11}}{m} & \frac{a_{12}}{m} \\ \frac{a_{21}}{m} & 1 + \frac{a_{22}}{m} \end{pmatrix}.$$

If terms $1/m^2$ and higher can be ignored then in the calculation of the determinant the non-diagonal terms can be ignored, and the determinant is simply the product of the diagonal elements:

$$\begin{aligned} Det(I + \frac{1}{m} A) &= (1 + \frac{a_{11}}{m})(1 + \frac{a_{22}}{m}) + \mathcal{O}(\frac{1}{m^2}) \\ &= 1 + \frac{1}{m}(a_{11} + a_{22}) + \mathcal{O}(\frac{1}{m^2}). \end{aligned}$$

In general one finds:

$$Det(I + \frac{1}{m} A) = 1 + \frac{1}{m} \mathrm{Tr}(A) + \mathcal{O}(\frac{1}{m^2}), \qquad \mathrm{Tr}(A) = \sum_k a_{kk}.$$

$\mathrm{Tr}(A)$ is called the trace of A, and is the sum of all elements along the diagonal. Inserting this in (A.6):

$$Det(e^A) = \lim_{m \to \infty} \left[1 + \frac{1}{m} \mathrm{Tr}(A) + \mathcal{O}(\frac{1}{m^2}) \right]^m = e^{\mathrm{Tr}(A)}. \qquad \text{(A.7)}$$

That terms of order $1/m^2$ can be ignored may be established easily.

Exercise A.2 Show that

$$\lim_{m \to \infty} (1 + \frac{y}{m^2})^m = \text{Constant}$$

by differentiation with respect to y. The constant is obviously one.