1.5 HW5

Excercise 1.5.2. Suppose all functions involved are well defined. For any square matrix A and $t \in \mathbb{R}$, we conclude that $\frac{\mathrm{d}}{\mathrm{d}t} f(tA) = f'(tA)A = Af'(tA)$.

WLOG we only prove this for Jordan block and $t = 1$ case. For $n \times n$ Jordan block $J(\lambda)$, we have

$$tJ(\lambda) = \begin{bmatrix} t\lambda & t & & \\ & \ddots & \ddots & \\ & & \ddots & t \\ & & & t\lambda \end{bmatrix} = BJ(t\lambda)B^{-1},$$

where $B = B(t) = \mathrm{diag}(t^{n-1}, \cdots, t, 1)$, and therefore

$$\begin{aligned} f(tJ(\lambda)) &= Bf(J(t\lambda))B^{-1} \\ &= \begin{bmatrix} f(t\lambda) & \frac{1}{1!}f'(t\lambda)t & \cdots & \frac{1}{(n-1)!}f^{(n-1)}(t\lambda)t^{n-1} \\ & \ddots & \ddots & \vdots \\ & & \ddots & \frac{1}{1!}f'(t\lambda)t \\ & & & f(t\lambda) \end{bmatrix} \\ &= \sum_{k=0}^{n-1} \frac{1}{k!} f^{(k)}(t\lambda)t^k N^k, \end{aligned}$$

where N is the $n \times n$ nilpotent matrix. Take the derivative to both sides, and we have (hereinafter J stands for $J(\lambda)$)

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t} f(tJ)\bigg|_{t=1} &= \sum_{k=0}^{n-1} \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{1}{k!} f^{(k)}(t\lambda)t^k N^k\right)\bigg|_{t=1} \\ &= \sum_{k=0}^{n-1} \lambda \cdot \frac{1}{k!} f^{(k+1)}(\lambda)N^k + \sum_{k=1}^{n-1} \frac{1}{(k-1)!} f^{(k)}(\lambda)N^k \\ &= \sum_{k=0}^{n-1} \lambda \cdot \frac{1}{k!} f^{(k+1)}(\lambda)N^k + \sum_{k=1}^{n} \frac{1}{(k-1)!} f^{(k)}(\lambda)N^k \qquad (N^n = O) \\ &= \sum_{k=0}^{n-1} \left(\lambda \cdot \frac{1}{k!} f^{(k+1)}(\lambda)N^k + \frac{1}{k!} f^{(k+1)}(\lambda)N^{k+1}\right). \end{aligned}$$

Note that $f'(J) = \sum_{k=0}^{n-1} \frac{1}{k!} f^{(k+1)}(\lambda)N^k$, thus

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t} f(tJ)\bigg|_{t=1} &= f'(J)(\lambda I + N) = f'(J)J \\ &= (\lambda I + N)f'(J) = Jf'(J). \end{aligned}$$

Q.E.D.

1. According to the conclusion above,

$$\frac{\mathrm{d}}{\mathrm{d}t} \sin(tA) = \cos(tA)A = A\cos(tA).$$

2. Let $M(t) = \begin{bmatrix} 2A & A \\ & (2+t)A \end{bmatrix}$, then $M(0) = \lim_{t\to 0} M(t)$, and

$$M(t) = XDX^{-1} = \begin{bmatrix} 1 & \frac{1}{t} \\ & 1 \end{bmatrix} \begin{bmatrix} 2A & \\ & (2+t)A \end{bmatrix} \begin{bmatrix} 1 & -\frac{1}{t} \\ & 1 \end{bmatrix}, \quad t \neq 0.$$

Therefore

$$f(M(t)) = Xf(D)X^{-1} = \begin{bmatrix} f(2A) & \frac{f((2+t)A)-f(2A)}{t} \\ & f((2+t)A) \end{bmatrix}, \quad t \neq 0.$$

Since $f(M(0)) = \lim_{t\to 0} f(M(t)) = \begin{bmatrix} f(2A) & B \\ & f(2A) \end{bmatrix}$, we have

$$B = \lim_{t\to 0} \frac{f((2+t)A) - f(2A)}{t} = \frac{\mathrm{d}}{\mathrm{d}t} f(tA)\bigg|_{t=2} = f'(2A)A.$$

3. A counter example is $f(x) = x^2$, and $A = \begin{bmatrix} 1 & \\ & -1 \end{bmatrix}$ and $B = \begin{bmatrix} 0 & 1 \\ & 0 \end{bmatrix}$. Therefore

$$f(A+tB) = \begin{bmatrix} 1 & t \\ & -1 \end{bmatrix}^2 = \begin{bmatrix} 1 & \\ & 1 \end{bmatrix},$$

and thus the derivative to it is zero matrix O, while $f'(A)B = 2AB = \begin{bmatrix} 0 & 2 \\ & 0 \end{bmatrix} \neq O$.