$\boxed{\textbf{Spring 2023} \quad \textbf{CIS 610}}$

Advanced Geometric Methods in Computer Science
Jean Gallier

Homework 1

January 20; Due February 9, 2023

Problem B1 (50). (a) Find two symmetric matrices, A and B, such that AB is not symmetric.

(b) Find two matrices A and B such that

$$e^A e^B \neq e^{A+B}.$$

Hint. Try

$$A = \pi \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix} \quad \text{and} \quad B = \pi \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix},$$

and use the Rodrigues formula.

(c) Find some square matrices A, B such that $AB \neq BA$, yet

$$e^A e^B = e^{A+B}.$$

Hint. Look for 2×2 matrices with zero trace.

Problem B2 (80 pts). Let $\mathrm{M}_n(\mathbb{C})$ denote the vector space of $n \times n$ matrices with complex coefficients (and $\mathrm{M}_n(\mathbb{R})$ denote the vector space of $n \times n$ matrices with real coefficients). For any matrix $A \in \mathrm{M}_n(\mathbb{C})$, let R_A and L_A be the maps from $\mathrm{M}_n(\mathbb{C})$ to itself defined so that

$$L_A(B) = AB, \quad R_A(B) = BA, \quad \text{for all } B \in \mathrm{M}_n(\mathbb{C}).$$

Check that L_A and R_A are linear, and that L_A and R_B commute for all A, B.

Let $\mathrm{ad_A} \colon \mathrm{M}_n(\mathbb{C}) \to \mathrm{M}_n(\mathbb{C})$ be the linear map given by

$$\mathrm{ad}_A(B) = L_A(B) - R_A(B) = AB - BA = [A, B], \quad \text{for all } B \in \mathrm{M}_n(\mathbb{C}).$$

Note that $[A, B]$ is the Lie bracket.

(1) Prove that if A is invertible, then L_A and R_A are invertible; in fact, $(L_A)^{-1} = L_{A^{-1}}$ and $(R_A)^{-1} = R_{A^{-1}}$. Prove that if $A = PBP^{-1}$ for some invertible matrix P, then

$$L_A = L_P \circ L_B \circ L_P^{-1}, \quad R_A = R_P^{-1} \circ R_B \circ R_P.$$

(2) Recall that the n^2 matrices E_{ij} defined such that all entries in E_{ij} are zero except the (i,j)th entry, which is equal to 1, form a basis of the vector space $\mathrm{M}_n(\mathbb{C})$. Consider the partial ordering of the E_{ij} defined such that for $i = 1, \ldots, n$, if $n \geq j > k \geq 1$, then then E_{ij} precedes E_{ik}, and for $j = 1, \ldots, n$, if $1 \leq i < h \leq n$, then E_{ij} precedes E_{hj}.

Draw the Hasse diagam of the partial order defined above when $n = 3$.

There are total orderings extending this partial ordering. How would you find them algorithmically? Check that the following is such a total order:

$$(1,3),\ (1,2),\ (1,1),\ (2,3),\ (2,2),\ (2,1),\ (3,3),\ (3,2),\ (3,1).$$

(3) Let the total order of the basis (E_{ij}) extending the partial ordering defined in (2) be given by

$$(i,j) < (h,k) \quad \text{iff} \quad \begin{cases} i = h \text{ and } j > k \\ \text{or } i < h. \end{cases}$$

Let $\lambda_1, \ldots, \lambda_n$ be the eigenvalues of A (not necessarily distinct). Using Schur's theorem, A is similar to an upper triangular matrix B, that is, $A = PBP^{-1}$ with B upper triangular, and we may assume that the diagonal entries of B in descending order are $\lambda_1, \ldots, \lambda_n$. If the E_{ij} are listed according to the above total order, prove that R_B is an upper triangular matrix whose diagonal entries are

$$(\underbrace{\lambda_n, \ldots, \lambda_1, \ldots, \lambda_n, \ldots, \lambda_1}_{n^2}),$$

and that L_B is an upper triangular matrix whose diagonal entries are

$$(\underbrace{\lambda_1, \ldots, \lambda_1}_{n} \ldots, \underbrace{\lambda_n, \ldots, \lambda_n}_{n}).$$

Hint. Figure out what are $R_B(E_{ij}) = E_{ij}B$ and $L_B(E_{ij}) = BE_{ij}$.

Use the fact that

$$L_A = L_P \circ L_B \circ L_P^{-1}, \quad R_A = R_P^{-1} \circ R_B \circ R_P,$$

to express $\mathrm{ad}_A = L_A - R_A$ in terms of $L_B - R_B$, and conclude tha t the eigenvalues of ad_A are $\lambda_i - \lambda_j$, for $i = 1, \ldots, n$, and for $j = n, \ldots, 1$.

(4) (**Extra Credit**) Let R be the $n \times n$ permutation matrix given by

$$R = \begin{pmatrix} 0 & 0 & \ldots & 0 & 1 \\ 0 & 0 & \ldots & 1 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 1 & \ldots & 0 & 0 \\ 1 & 0 & \ldots & 0 & 0 \end{pmatrix}.$$

Observe that $R^{-1} = R$. I checked for $n = 3$ that in the basis (E_{ij}) ordered as above, the matrix of L_A is given by $A \otimes I_3$, and the matrix of R_A is given by $I_3 \otimes RA^\top R$. Here, $\otimes$ the *Kronecker product* (also called *tensor product*) of matrices.

Given an $m \times n$ matrix $A = (a_{ij})$ and a $p \times q$ matrix $B = (b_{ij})$, the *Kronecker product* (or *tensor product*) $A \otimes B$ of A and B is the $mp \times nq$ matrix

$$A \otimes B = \begin{pmatrix} a_{11}B & a_{12}B & \cdots & a_{1n}B \\ a_{21}B & a_{22}B & \cdots & a_{2n}B \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1}B & a_{m2}B & \cdots & a_{mn}B \end{pmatrix}.$$

It can be shown (and you may use these facts without proof) that $\otimes$ is associative and that

$$\begin{aligned} (A \otimes B)(C \otimes D) &= AC \otimes BD \\ (A \otimes B)^\top &= A^\top \otimes B^\top, \end{aligned}$$

whenever AC and BD are well defined.

Prove that for any $n \geq 1$, the matrix of L_A is given by $A \otimes I_n$, and the matrix of R_A is given by $I_n \otimes RA^\top R$. Use this result to give another proof of the fact that the eigenvalues of ad_A are $\lambda_i - \lambda_j$, for $i = 1, \ldots, n$, and for $j = n, \ldots, 1$.

Note that if instead of the ordering

$$E_{1n}, E_{1n-1}, \ldots, E_{11}, E_{2,n}, \ldots, E_{21}, \ldots, E_{nn}, \ldots, E_{n1},$$

that I proposed you use the standard lexicographic ordering

$$E_{11}, E_{12}, \ldots, E_{1n}, E_{21}, \ldots, E_{2n}, \ldots, E_{n1}, \ldots, E_{nn},$$

then the matrix representing L_A is still $A \otimes I_n$, but the matrix representing R_A is $I_n \otimes A^\top$. In this case, if A is upper-triangular, then the matrix of R_A is *lower triangular*. This is the motivation for using the first basis (avoid upper becoming lower).

Problem B3 (80 pts). Given any two matrices $A, X \in \mathrm{M}_n(\mathbb{C})$, define the function $f_A \colon \mathrm{M}_n(\mathbb{C}) \to \mathrm{M}_n(\mathbb{C})$ by

$$f_A(X) = \sum_{p,q \geq 0} \frac{A^p X A^q}{(p+q+1)!}.$$

(1) Prove that

$$f_A = \sum_{p,q \geq 0} \frac{L_A^p \circ R_A^q}{(p+q+1)!}.$$

(2) Prove that

$$\mathrm{ad}_A \circ f_A = \sum_{k \geq 1} \frac{1}{k!}(L_A^k - R_A^k).$$

(3) Prove that

$$\mathrm{ad}_A \circ f_A = e^{L_A} \circ (\mathrm{id} - e^{-\mathrm{ad}_A}).$$

Check that

$$e^{L_A}(B) = e^A B.$$

Conclude that

$$\mathrm{ad}_A \circ f_A = e^A(\mathrm{id} - e^{-\mathrm{ad}_A}).$$

Observe that

$$\mathrm{id} - e^{-\mathrm{ad}_A} = \sum_{k=0}^{\infty} \frac{(-1)^k \mathrm{ad}_A^{k+1}}{(k+1)!}$$

so it would be tempting to say that

$$f_A = e^A \sum_{k=0}^{\infty} \frac{(-1)^k \mathrm{ad}_A^k}{(k+1)!},$$

but I don't know a simple way of justifying this fact!

(4) Prove that

$$d(\exp)_A(X) = \sum_{p,q\geq 0} \frac{A^p X A^q}{(p+q+1)!} = f_A(X).$$

Remark: It is known that

$$d(\exp)_A = e^A \sum_{k=0}^{\infty} \frac{(-1)^k \mathrm{ad}_A^k}{(k+1)!},$$

so the bold unsubstantiated conclusion in (3) is actually correct. In fact, it is customary to use the notation

$$\frac{\mathrm{id} - e^{-\mathrm{ad}_A}}{\mathrm{ad}_A}$$

for the power series

$$\sum_{k=0}^{\infty} \frac{(-1)^k \mathrm{ad}_A^k}{(k+1)!},$$

and the formula for the derivative of exp is usually stated as

$$d(\exp)_A = e^A \left(\frac{\mathrm{id} - e^{-\mathrm{ad}_A}}{\mathrm{ad}_A} \right).$$

Problem B4 (20 pts). Let $f\colon \mathbb{R}^2 \to \mathbb{R}$ be the function given by

$$f(x,y) = \begin{cases} \frac{x^2 y}{x^4+y^2} & \text{if } (x,y) \neq (0,0) \\ 0 & \text{if } (x,y) = (0,0). \end{cases}$$

(a) Compute the directional derivative $\mathrm{D}_u f(0,0)$ of f at $(0,0)$ for every vector $u = (u_1, u_2) \neq 0$.

(b) Prove that the derivative $\mathrm{D}f(0,0)$ does not exist. What is the behavior of the function f on the parabola $y = x^2$ near the origin $(0,0)$?

Problem B5 (40 pts). (a) Let $f\colon \mathrm{M}_n(\mathbb{R}) \to \mathrm{M}_n(\mathbb{R})$ be the function defined on $n \times n$ matrices by

$$f(A) = A^2.$$

Prove that

$$\mathrm{D}f_A(H) = AH + HA,$$

for all $A, H \in \mathrm{M}_n(\mathbb{R})$.

(b) Let $f\colon \mathrm{M}_n(\mathbb{R}) \to \mathrm{M}_n(\mathbb{R})$ be the function defined on $n \times n$ matrices by

$$f(A) = A^3.$$

Prove that

$$\mathrm{D}f_A(H) = A^2H + AHA + HA^2,$$

for all $A, H \in \mathrm{M}_n(\mathbb{R})$.

TOTAL: 350 **points.**