Notes on Matrix Exponentiation etc.

Define $\quad A \quad n \times n$ real matrix (or $\mathbb{C}$-valued entries OK too)

$$e^A = I + A + \frac{1}{2}A^2 + \frac{1}{3!}A^3 + \frac{1}{4!}A^4 \cdots$$

Series converges: use "operator norm"

$$\|A\|_{op} = \max_{\vec{x} \in \mathbb{R}^n \ \|\vec{x}\|=1} \|A\vec{x}\|$$

and $\quad \|AB\|_{op} \leq \|A\|_{op} \|B\|_{op}$

$$\|A+B\|_{op} \leq \|A\|_{op} + \|B\|_{op}$$

(as in homework)

Note: If $AB = BA$ then $\quad e^{A+B} = e^A e^B$

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

Reason $\quad e^x e^y = e^{x+y}$ $\quad , x, y \in \mathbb{R}$

$\Rightarrow$ (series for e^x)(series for e^y) = series for e^{x+y}

in formal multiplication terms

e.g. $\quad \left(1 + x + \frac{x^2}{2!} + \cdots\right)\left(1 + y + \frac{y^2}{2!} \cdots\right)$

$$= 1 + (x+y) + \frac{1}{2}(x^2 + 2xy + y^2) + \cdots$$

$$= e^{x+y} \text{ formally}$$

General situation:

$$\sum_{n+m=N} \frac{1}{n!}\frac{1}{m!} x^n y^m$$

$$= \sum_{n=0}^{N} \frac{1}{N!} \frac{N!}{n!(N-n)!} x^n y^{N-n}$$

$$= \frac{1}{N!}\left(\sum_{n+m=N} \binom{N}{n} x^n y^{N-n}\right)$$

$$= \frac{1}{N!}(x+y)^N$$

(here $0! = 1$ by convention)

So $e^A e^B = e^{A+B}$ works if A, B are $n \times n$ matrices which commute ($AB = BA$)

Natural question:
What does it take to make e^A orthogonal (and hence in $SO(n)$ since $\det e^A = e^{\operatorname{tr} A} > 0$; see homework).

A skew symmetric works; $A^T = -A$ (T = transpose)
Since then $AA^T = A^T A = -A^2$)
so $e^A (e^A)^T = e^A e^{A^T} = e^A e^{-A} = e^0 = I$.

Now it is ~~not~~ true in general that $e^A = I \Rightarrow A$ skew sym

But what is true is that

$e^{tA} \in SO(n)$ for all t with $|t| < \varepsilon$ some $\varepsilon > 0$

$\Rightarrow$ A skewsymmetric.

Reason: $I = e^{tA}(e^{tA})^T$

$= (I + tA + t^2 \text{ or higher deg.})(I + tA^T + t^2 \text{ or higher degree})$

$= I + t(A + A^T) + (t^2 \text{ or higher degree})$

all t with $|t| < \varepsilon \Rightarrow A + A^T = 0$ or $A^T = -A$.

Similar analyses:

$\det e^{tA} = 1$ all t with $|t| < \varepsilon$, some $\varepsilon > 0$

$\Leftrightarrow$ $\operatorname{tr} A = 0$

(Actually $\det e^B = 1 \Leftrightarrow \operatorname{tr} B = 0$ since $\det e^B = e^{\operatorname{tr} B}$: see homework)

Now $A \xrightarrow{\operatorname{Exp}} e^A$ is differentiable from $\mathbb{R}^{n^2}$ to $\mathbb{R}^{n^2}$ at $A = I$ and its differential $d\operatorname{Exp}|_{0 \text{ matrix}}$ is I since

$A \to I + A + \frac{A^2}{2!} + \cdots$

So by Inverse Function Theorem,

$SO(n)$ near I looks like / is differentiably equivalent to) a neighborhood of $\vec{0}$ matrix in skew symmetric matrices. (which is $\mathbb{R}^{n(n-1)/2}$).

So $SO(n)$ is like a $\frac{n(n-1)}{2}$ dim "surface"

in $\mathbb{R}^{n^2}$ near I.

Think about what it looks like at some other point A_0. Still looks same

Since $A \to A_0 A$ is a differentiable invertible function on $A \in \mathbb{R}^{n^2}$

$SO(n)$ is a "differentiable manifold" of dimension $\frac{n(n-1)}{2}$; looks around each point like an open set in $\mathbb{R}^{n(n-1)/2}$

Def: A Lie group is a group that is also a differentiable manifold with the group operations differentiable.

So $Gl(n, \mathbb{R})$ (invertible $n \times n$ matrices)

$SO(n,\mathbb{R})$

$SL(n,\mathbb{R})$ = matrices with $\det = 1$

are Lie groups.

Big subject: Lie groups.

Fundamental approach: G a Lie group

e = identity, T_eG = "tangent vectors" at e. $\overset{\text{def}}{=}$ Lie algebra of G.

The tangent space (= set of all tangent vectors) at e turns out to have an algebra structure itself.

In matrix case, $A, B \in$ tangent space

$\Rightarrow$ $AB - BA \in$ tangent space

Example:

A, B skew symmetric (tangent space at e of $SO(n)$)

$\Rightarrow$ $AB - BA$ skew symmetric

$$(AB - BA)^T = (AB)^T - (BA)^T$$
$$= B^TA^T - A^TB^T$$

$= (-B)(-A) - (-A)(-B)$

$= BA - AB = -(AB - BA)$ ✓

Notation: $AB - BA = [A, B]$ Lie bracket

$[\,,\,]$ is a binary operation, but it is

not associative! in general.

$[[A, B], C] \overset{?}{=} [A, [B, C]]$

$$[[A,B],C] = (AB - BA)C - C(AB - BA) = \cancel{ABC} - BAC - CAB - \cancel{CBA}$$

$$[A,[B,C]] = A(BC - CB) - (BC - CB)A = ABC - ACB - ABC - \cancel{CBA}$$

$BAC + CAB = ACB + ABC$

might not work!

For instance if $A = B$

$\cancel{A^2C} + CA^2 = ACA + \cancel{A^2C}$

$CA^2 = ACA$ might not be true;

(if A inv)

$\Downarrow$

$CA = AC$

But $[\,,\,]$ does satisfy some

kind of identity. The "Jacobi identity")

$$[A,[B,C]]+[C,[A,B]]+[B,[C,A]]=0$$

(exercise!)

(sum over cyclic permutation!)

This sort of substitutes for associativity in the algebraic analysis.

Big idea: the Lie algebra encodes all the multiplication information of the Lie group.

Isomorphic Lie algebras

$\Longleftrightarrow$ local isomorphism of Lie groups.

Something for the future!

Back to $SO(n)$ and $Gl(n,\mathbb{R})$ etc.

Given a matrix A close to I_n, is there a skewsymmetric B such that

$$e^B = A.$$

Yes. This works if $\|A-I\|_{op}<1$.

Reason for numbers

$$e^{x-\frac{x^2}{2}+\frac{x^3}{3}-\frac{x^4}{4}\cdots} = 1+x \text{ if } |x|<1$$

(Proof: $\frac{1}{1+x} = 1 - x + x^2 - x^3 \cdots$

integrate term by term to get

$\ln x = x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} \cdots$)

As before this works for matrices too

So if $I + X = A$ $\quad \|X\|_{op} < 1$

then $e^{X - \frac{X^2}{2} + \frac{X^3}{3} \cdots} = I + X = A.$

Matrices "near I have "matrix logs".

Now what about $SO(n)$?

In this case, if $A \in SO(n)$ $\det A = 1$

skew symmetric $B \ni e^B = A$

(you do not need A close to I.)

Reason: According to linear algebra

A in some orthogonal basis has the form of 2×2 blocks of det 1 (plus if A is odd x odd a $+1$ axis).

But each 2×2 block

$$\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} = e^{B_j}$$

$B_j = \begin{pmatrix} 0 & -\theta \\ \theta & 0 \end{pmatrix}$ $\qquad$ Exercise: Compute this works

So one gets a log for A, skew sym in 2×2 block form (plus a 0 1×1 block in the odd case)

Note that this all fits together

$B \xrightarrow{\mathrm{Exp}} e^B$ is differentiable with nonsingular differential at 0 (image $= I$)

skew symmetric matrices are $\frac{n(n-1)}{2}$ dim and orthogonal matrices are $\frac{n(n-1)}{2}$ dim (homework). So it makes sense that Exp maps skew sym. differentially one-to-one onto (near 0, $\mathrm{Exp}(0) = I$) with diff. inverse. It makes sense but the proof needs that 2×2 block decomposition (or something else additional to dimension counting).