Eigenvalues of $\hat{L}_{f,X}$

Recap: Fréchet derivative of f in $X \in \mathbb{C}^{n\times n}$: is the linear operator s.t.

$$f(X+E) = f(X) + L_{f,X}[E] + o(\|E\|)$$

ex: $f(X) = X^2$ $\quad L_{f,X}[E] \to XE + EX$

$$\hat{L}_{f,X} = X^T \otimes I + I \otimes X$$

Th: Let $X \in \mathbb{C}^{n\times n}$ have eigenvalues $\lambda_1, \lambda_2, \dots \lambda_n$ (with multiplicity). Then, the n^2 eigenvalues of $\hat{L}_{f,X}$ are

$$f[\lambda_i, \lambda_j] = \begin{cases} \dfrac{f(\lambda_i) - f(\lambda_j)}{\lambda_i - \lambda_j} & \text{if } \lambda_i \neq \lambda_j \\ f'(\lambda_i) & \text{if } \lambda_i = \lambda_j \end{cases}$$

for $i, j = 1, 2, \dots n$

Prf: We can replace f with a polynomial $p(x)$ such that $f^{(k)}(\lambda_i) = p^{(k)}(\lambda_i)$ for k up to $2n-1$ so that

$$f\left(\begin{bmatrix} X & E \\ 0 & X \end{bmatrix}\right) = p\left(\begin{bmatrix} X & E \\ 0 & X \end{bmatrix}\right)$$

and hence $L_{f,X}[E] = L_{p,X}[E]$ (it is the (1,2) block of $\cdot$)

If $p(x) = c_0 + c_1 x + c_2 x^2 + \dots + c_d x^d$

$$p(X+E) = c_0 I + c_1 (X+E) + c_2 (X+E)^2 + \dots + c_d (X+E)^d$$

$$\begin{aligned} &= c_0 I + c_1 X + c_1 E + c_2 X^2 + c_2 (XE+EX) + c_2 E^2 \\ &\quad + c_3 X^3 + c_3 (EX^2 + XEX + X^2 E) + o(\|E\|) \\ &\quad + \dots + c_d X^d + c_d (EX^{d-1} + XEX^{d-2} + X^2 E X^{d-3} + \dots X^{d-1} E) \\ &\quad + o(\|E\|) \end{aligned}$$

$$= p(X) + \underbrace{\sum_{k=1}^{d} c_k \sum_{h=1}^{k} X^{k-h} E X^{h-1}}_{L_{f,X}[E]} + o(\|E\|)$$

$$\hat{L}_{f,X} = \sum_{k=1}^{d} c_k \sum_{h=1}^{k} (X^{h-1})^T \otimes X^{k-h}$$

Take Schur forms $X = Q_1 U_1 Q_1^*$ $\quad X^T = Q_2 U_2 Q_2^*$

Then $X^{k-h} = Q_1 U_1^{k-h} Q_1^*$ $\quad (X^{h-1})^T = Q_2 U_2^{h-1} Q_2^*$

and $\hat{L}_{f,X} = \sum_k c_k \sum_h Q_2 U_2^{h-1} Q_2^* \otimes Q_1 U_1^{k-h} Q_1^*$

$$= \underbrace{(Q_2 \otimes Q_1)}_{\text{orth}} \underbrace{\left(\sum_k c_k \sum_h U_2^{h-1} \otimes U_1^{k-h}\right)}_{\text{triangular} := U} \underbrace{(Q_2 \otimes Q_1)^*}_{\text{orth}^*}$$

This is a Schur decomposition, we can read off the eigenvalues from the diagonal of U

$$U_{i+n(j-1),\, i+n(j-1)} = \sum_{k=1}^{d} \sum_{h=1}^{k} c_k \lambda_i^{h-1} \lambda_j^{k-h} .$$

$$= \sum_{k=1}^{d} c_k \sum_{h=1}^{k} \lambda_i^{h-1} \lambda_j^{k-h} = \sum_{k=1}^{d} c_k \left(\lambda_j^{k-1} + \lambda_i \lambda_j^{k-2} + \lambda_i^2 \lambda_j^{k-3} + \dots + \lambda_i^{k-1} \right)$$

if $\lambda_i \neq \lambda_j$: $$= \sum_{k=1}^{d} c_k \frac{\lambda_i^k - \lambda_j^k}{\lambda_i - \lambda_j} = \frac{p(\lambda_i) - p(\lambda_j)}{\lambda_i - \lambda_j} = p[\lambda_i, \lambda_j]$$

if $\lambda_i = \lambda_j$ $$= \sum_{k=1}^{d} c_k k \lambda_i^{k-1} = p'(\lambda_i).$$

Ex: $f(x) = \sqrt{x}$, principal square root

Eigenvalues of $\hat{L}_{f,X}$ are $$\begin{cases} \dfrac{\lambda_i^{1/2} - \lambda_j^{1/2}}{\lambda_i - \lambda_j} & \lambda_i \neq \lambda_j \\ \dfrac{1}{2\lambda_i^{1/2}} & \lambda_i = \lambda_j \end{cases} \quad i,j = 1, \dots, n$$

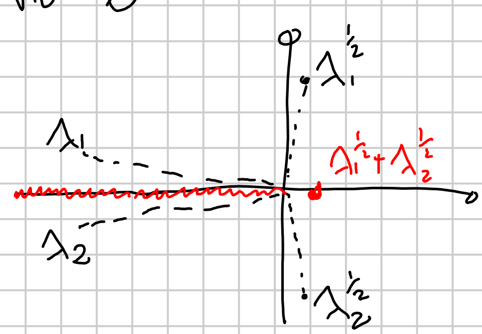
For which choices of λ_i, λ_j are these eigenvalues large in modulus?

Ex: $n=2$ $\quad \Lambda(X) = \{\lambda_1, \lambda_2\}$ $\quad \lambda_1 \neq \lambda_2$

$$\Lambda(\hat{L}_{f,X}) = \left\{ \frac{1}{2\lambda_1^{1/2}}, \frac{\lambda_1^{1/2} - \lambda_2^{1/2}}{\lambda_1 - \lambda_2}, \underbrace{\frac{\lambda_1^{1/2} - \lambda_2^{1/2}}{\lambda_1 - \lambda_2}}_{= \frac{1}{\lambda_1^{1/2} + \lambda_2^{1/2}}}, \frac{1}{2\lambda_2^{1/2}} \right\}$$

Large eigenvalues if:

- X has an eigenvalue close to 0
- X has eigenvalues close to each other, but on opposite sides of the negative real axis. (branch cut line)



More in general, for any function f, large eigenvalues come from:

- points with large $f'(\lambda_i)$
- pairs of points close to a discontinuity in f: $\lambda_i - \lambda_j$ small, but $f(\lambda_i) - f(\lambda_j)$ not small.

This gives insight on cases in which to expect ill-conditioning.

If X is normal, U_1 and U_2 can be taken diagonal, so $L_{f,X} = (Q_1 \otimes Q_2)(\ \cdots\cdots\ )(Q_1 \otimes Q_2)^*$ is an orthogonal eigendecomposition, and this is the whole story.

If X is not normal, $\|\hat{L}_{f,X}\| = \sigma_{max}(\hat{L}_{f,X}) \geq |\lambda_{max}(\hat{L}_{f,X})|$.

We can at least give a bound: if X is diagonalizable, $X = V \operatorname{diag}(\lambda_1, \lambda_2, \dots \lambda_n) V^{-1}$, we can repeat this argument with this decomposition, and obtain

$$\hat{L}_{f,X} = (V^T \otimes V) \cdot \operatorname{diag}\left(f[\lambda_i, \lambda_j],\ i,j = 1,\dots n\right) \cdot (V^T \otimes V)^{-1}$$

$$\|\hat{L}_{f,X}\|_2 \leq \kappa_2^2(V) \max_{i,j} \left| f[\lambda_i, \lambda_j] \right|$$

$\|V^T \otimes V\|_2 = \|V^T\|_2 \cdot \|V\|_2$
$= \|V\|_2^2$

For non-normal matrices, a third cause of ill-conditioning is:

- $\kappa_2(V)$ large.

Numerical methods for matrix functions

Based on:

- decompositions $X = VDV^{-1}$, $X = QUQ^*$
- interpolation / approximation: replace f with a polynomial or rational function
- Cauchy integral formula $f(X) = \frac{1}{2\pi i} \int_\Gamma f(z)(zI - X)^{-1} dz$
- Special tricks for certain functions, $\exp(2a) = \exp(a)^2$ or root finding iterations, ex. $X^{\frac{1}{2}}$ via Newton on $X^2 - A = 0$.
- Arnoldi iteration

Simplest method: if X diagonalizable, $X = VDV^{-1}$

$$f(X) = f(VDV^{-1}) = V f(D) V^{-1} = V \cdot \begin{bmatrix} f(\lambda_1) & & & \\ & f(\lambda_2) & & \\ & & \ddots & \\ & & & f(\lambda_n) \end{bmatrix} V^{-1}$$

If X Hermitian / symmetric / normal, V is orthogonal, and this method is stable, and a good one.

Otherwise errors may be amplified:

suppose you make an error in computing $f(\lambda_i) = y_i$, leading to $\tilde{y}_i$ with $|\tilde{y}_i - y_i| \leq \varepsilon$

Then, even ignoring other sources of numerical errors, the computed $\tilde{Y} = V \begin{bmatrix} \tilde{y}_1 & & & \\ & \tilde{y}_2 & & \\ & & \ddots & \\ & & & \tilde{y}_n \end{bmatrix} V^{-1}$

$$\|\tilde{Y} - f(X)\| = \left\| V \begin{bmatrix} \tilde{y}_1 & & \\ & \ddots & \\ & & \tilde{y}_n \end{bmatrix} V^{-1} - V \begin{bmatrix} y_1 & & \\ & \ddots & \\ & & y_n \end{bmatrix} V^{-1} \right\| =$$

$$= \left\| V \begin{bmatrix} \tilde{y}_1 - y_1 & & \\ & \ddots & \\ & & \tilde{y}_n - y_n \end{bmatrix} V^{-1} \right\| \leq \|V\| \cdot \varepsilon \cdot \|V^{-1}\| = \varepsilon \cdot \kappa(V)$$

If $\kappa(V)$ is large, we can expect trouble.

Ex:

```
>> A = [3 -1; 1 1+1e-15]
A =
    3.0000   -1.0000
    1.0000    1.0000
>> [V, D] = eig(A)
V =
   0.7071 + 0.0000i   0.7071 + 0.0000i
   0.7071 - 0.0000i   0.7071 + 0.0000i
D =
   2.0000 + 0.0000i   0.0000 + 0.0000i
   0.0000 + 0.0000i   2.0000 - 0.0000i
>> cond(V)
ans =
   6.7109e+07
>> Y = V*sqrt(D)/V;
>> norm(Y^2-A)
ans =
   1.8531e-08
>> norm(Y^2-A) / norm(A)
ans =
   5.7263e-09
```

Instead, we try to replace $f(x)$ with the polynomial such that $p(2)=f(2)$, $p'(2)=f'(2)$, $\deg p = 1$

$$p(x) = \frac{1}{\sqrt{8}}x + \frac{1}{\sqrt{2}}$$

$$Z = p(A) \qquad \|Z^2 - A\| / \|A\| = O(10^{-16})$$

However, $p(A) \neq f(A)$ in exact arithmetic, because $p(x)$ is not the interpolating polynomial in $\Lambda(A) = \{2.00.., 2.00...\}$

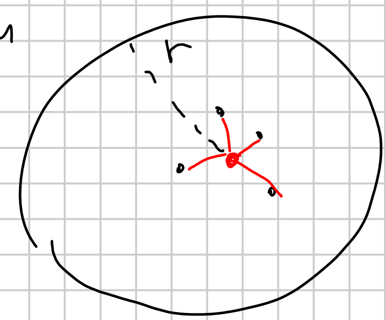
The correct thing to do would have been replacing $f(x)$ not with a polynomial but with its Taylor series

$$f(x) = \underbrace{f(2) + f'(2)(x-2)}_{p(x)} + \frac{f''(2)}{2}\underbrace{(x-2)^2}_{O(10^{-14}) \text{ on my example matrix}} + \dots$$

This suggests a general method to approximate matrix functions:

1. Choose a center for Taylor expansion
   Cheap but effective choice:

$$\alpha = \frac{\lambda_1 + \lambda_2 + \dots + \lambda_n}{n} = \frac{\operatorname{Tr}(A)}{n}$$



2. Compute $f(A) \approx f(\alpha)I + f'(\alpha)(A - \alpha I) + \frac{f''(\alpha)}{2}(A - \alpha I)^2 + \dots$
   stopping when the method converges (with a suitable stopping criterion).

<u>Theorem</u> (convergence of Taylor series)

Suppose $f(x) = \sum_{k=0}^{\infty} \frac{f^{(k)}(\alpha)}{k!}(x-\alpha)^k$ is a Taylor expansion with convergence radius $r>0$.

Then,
$$\lim_{d\to\infty} \sum_{k=0}^{d} \frac{f^{(k)}(\alpha)}{k!}(x-\alpha)^k = f(A)$$
for all matrices A with $\Lambda(A) \subseteq \{\lambda : |\lambda-\alpha| < r\}$.

<u>Proof</u>: We can reduce to the case of Jordan blocks

Indeed, let us set $p_d(x) = \sum_{k=0}^{d} \frac{f^{(k)}(\alpha)}{k!}(x-\alpha)^k$

Then, $A = VJV^{-1}$
$$\lim_{d\to\infty} p_d(A) = \lim_{d\to\infty} p_d(VJV^{-1}) = V \lim_{d\to\infty} \begin{bmatrix} p_d(J_1) & & \\ & \ddots & \\ & & p_d(J_s) \end{bmatrix} V^{-1}$$

We can conclude if we prove convergence on the Jordan blocks.

$$p_d(J) = \begin{bmatrix} p_d(\lambda) & p_d'(\lambda) & \cdots & \frac{1}{(k-1)!} p_d^{(k-1)}(\lambda) \\ & \ddots & \ddots & \vdots \\ & & \ddots & p_d'(\lambda) \\ & & & p_d(\lambda) \end{bmatrix}$$

On diagonal elements, $\lim_{d\to\infty} p_d(\lambda) = f(\lambda)$ since λ is in the radius of convergence

In the strictly upper triangular part,

$$\lim_{d\to\infty} \underbrace{\frac{1}{j!} P_d^{(j)}(\lambda)}$$

power series for $f^{(j)}(x)$, truncated to degree $d-j$

The power series for $f^{(j)}(x)$ has the same radius of convergence as the series for $f(x)$, which is a general fact on power series seen in analysis courses.

So $\lim_{d\to\infty} \frac{1}{j!} P_d^{(j)}(\lambda) = \frac{1}{j!} f^{(j)}(\lambda)$.