Conditioning of computing matrix functions

Recall: the absolute condition number of a differentiable $f : \mathbb{R}^m \to \mathbb{R}^n$ is the norm of its Jacobian.

$f(\tilde{x}) = f(x+h) = f(x) + \nabla_x f \cdot h + o(h)$ implies

$$\kappa_{abs}(f,x) = \lim_{\varepsilon \to 0} \sup_{\|\tilde{x}-x\| \le \varepsilon} \frac{\|f(\tilde{x}) - f(x)\|}{\|\tilde{x} - x\|} = \|\nabla f\|$$

$$\kappa_{rel}(f,x) = \lim_{\varepsilon \to 0} \sup_{\frac{\|\tilde{x}-x\|}{\|x\|} \le \varepsilon} \frac{\frac{\|f(\tilde{x})-f(x)\|}{\|f(x)\|}}{\frac{\|\tilde{x}-x\|}{\|x\|}} = \kappa_{abs}(f,x)\frac{\|x\|}{\|f(x)\|}.$$

Fréchet derivative

The Fréchet derivative is an "operator version" of the Jacobian.

Definition

The Fréchet derivative of a matrix function f is the linear operator $L_{f,X} : \mathbb{R}^{m\times m} \to \mathbb{R}^{m\times m}$ (when it exists) such that

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

I.e., in a neighbourhood of X, f behaves like a linear function.

Example

$f(x) = x^2$, $f(X) = X^2$.

$$(X+E)^2 = X^2 + XE + EX + E^2 = X^2 + \underbrace{XE + EX}_{L_{f,X}(E)} + o(\|E\|^2).$$

$L_{f,X}$ is a linear operator that maps matrices to matrices — we can consider its vectorized version:

$$\widehat{L} : \operatorname{vec} E \mapsto \operatorname{vec} L_{f,X}(E).$$

In this case,

$$\widehat{L} = X^T \otimes I + I \otimes X.$$

$\widehat{L}$ (an $n^2 \times n^2$ matrix) is the "usual" Jacobian of the map $\operatorname{vec} X \mapsto \operatorname{vec} f(X)$.

Properties

Follow from those of Jacobians:

- $L_{f+g,X} = L_{f,X} + L_{g,X}$.
- $L_{f \circ g,X} = L_{f,g(X)} \circ L_{g,X}$.
- $L_{f^{-1},f(X)} = L_{f,X}^{-1}$.

Example Let $g(y) = \sqrt{y}$ (principal branch: we take the root in the right half-plane), Y with no real nonpositive eigenvalue.
Then $g(y)$ is the inverse of $f(x) = x^2$, and its Fréchet derivative $F = L_{g,Y}(E)$ is the matrix such that $L_{f,X}(F) = E$, i.e.,

$$XF + FX = E, \quad X = f(Y) = Y^{1/2}.$$

(solution of a Sylvester equation). X has eigenvalues in the right half-plane, so the Sylvester equation is always solvable: $\Lambda(X) \cap \Lambda(-X) = \emptyset$.

Derivative of the exponential

Derivative of the matrix exponential:

$$\begin{aligned}\exp(X+E) &= I + (X+E) + \frac{1}{2}(X+E)^2 + \frac{1}{3!}(X+E)^3 + \dots \\ &= I + (X+E) + \frac{1}{2}(X^2 + EX + XE + E^2) + \frac{1}{3!}(X^3 + \dots) \\ &= \exp(X) + E + \frac{1}{2}(EX + XE) + \frac{1}{3!}(X^2E + XEX + X^2E) \\ &\quad + \dots + O(\|E\|^2)\end{aligned}$$

The series converges, but there is no easy closed form.

$$\widehat{L} = I + \frac{1}{2}(I \otimes X + X^T \otimes I) + \frac{1}{3!}(I \otimes X^2 + X^T \otimes X + (X^2)^T \otimes I) + \dots$$

Trick to compute $L_{f,X}(E)$

Let f be Fréchet differentiable. Then,

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

Proof (sketch) Evaluate $f\left(\begin{bmatrix} X+\varepsilon E & E \\ 0 & X \end{bmatrix}\right)$ by block-diagonalizing.

We need $\begin{bmatrix} I & Z \\ 0 & I \end{bmatrix}$, where Z solves $(X+\varepsilon E)Z - ZX = E$, which has solution $Z = -\frac{1}{\varepsilon} I$ (to block-diagonalize it, it is sufficient to find one solution, even if the Sylvester equation is singular). The evaluation gives $\begin{bmatrix} f(X+\varepsilon E) & \frac{f(A+\varepsilon E)-f(X)}{\varepsilon} \\ 0 & f(X) \end{bmatrix}$.

Existence of the Fréchet derivative

Theorem

If $f \in \mathcal{C}^{2m-1}(U)$, then $L_{f,X}$ exists for each $X \in \mathbb{R}^{m \times m}$ with eigenvalues in U.

Proof (sketch) The proof of the previous theorem shows that the directional derivatives of f (seen as a map $\mathbb{R}^{m^2} \to \mathbb{R}^{m^2}$) exist and are continuous (since matrix functions are continuous). It is a classical result in multivariate calculus that then f is continuously differentiable.

Fréchet derivative and condition number

Hence, $\kappa_{abs}(f, X) = \|L_{f,X}\|$.

...with some attention to what 'norm' means here.

The norm used for $\|\widetilde{X} - X\|$ is any matrix norm on $n \times n$ matrices, and $\|L_{f,X}\|$ is the 'operator norm' (on $n^2 \times n^2$ matrices) induced by it.

Easy case If we take $\|\widetilde{X} - X\|_F$, it corresponds to $\|\operatorname{vec} X\|_2$, so $\kappa_{abs}(f, X) = \|\widehat{L}_{f,X}\|_2$.

Harder cases For all other norms ($\|\widetilde{X} - X\|_2$ in particular), no equivalent simple expression for the 'induced operator norm'.

Eigenvalues of Fréchet derivatives [Higham book '08, Ch. 3]

Theorem

Let X have eigenvalues $\lambda_1, \dots, \lambda_n$. The eigenvalues of $L_{f,X}$ are

$$f[\lambda_i, \lambda_j] := \begin{cases} \frac{f(\lambda_i)-f(\lambda_j)}{\lambda_i-\lambda_j} & i \neq j, \\ f'(\lambda_i) & i = j. \end{cases}$$

Proof First of all, replace $f(x)$ with its interpolating polynomial $p(x)$ on the spectrum of A (and twice the multiplicities, to make sure the derivatives exist: $\left[\begin{smallmatrix} X & E \\ 0 & X \end{smallmatrix}\right]$ must be well-defined).

(continues)

$$
\begin{aligned}
p(X+E) &= p_0 + (X+E) + p_1(X+E)^2 + p_2(X+E)^3 + \dots \\
&= p_0 + p_1(X+E) + p_2(X^2 + EX + XE + E^2) + p_3(X^3 + \dots) \\
&= p(X) + p_1 E + p_2(EX + XE) + p_3(X^2 E + XEX + X^2 E) \\
&\qquad + \cdots + O(\|E\|^2)
\end{aligned}
$$

Vectorizing,

$$
\widehat{L}_{f,X} = p_1 I + p_2(I \otimes X + X^T \otimes I) + p_3(I \otimes X^2 + X^T \otimes X + (X^2)^T \otimes I) + \dots
$$

i.e.,

$$
\hat{L}_{f,X} = \sum_{k=0}^{d} p_k \sum_{h=1}^{k} (X^{k-h})^T \otimes X^{h-1}
$$

Eigenvalues of Fréchet derivatives

$$\hat{L}_{f,X} = \sum_{k=0}^{d} p_k \sum_{h=1}^{k} (X^{k-h})^T \otimes X^{h-1}$$

Take Schur forms $X = Q_1 T_1 Q_1^T$, $X^T = Q_2 T_2 Q_2^T$ to obtain a triangular matrix T.

On its diagonal, we can read off the eigenvalues

$$\begin{aligned} T_{i+n(j-1),i+n(j-1)} &= \sum_{k=0}^{d} p_k \Big(\sum_{h=1}^{k} \lambda_i^{k-h} \lambda_j^{h-1}\Big) = \sum_{k=0}^{d} p_k \frac{\lambda_i^k - \lambda_j^k}{\lambda_i - \lambda_j} \\ &= \frac{p(\lambda_i) - p(\lambda_j)}{\lambda_i - \lambda_j} = \frac{f(\lambda_i) - f(\lambda_j)}{\lambda_i - \lambda_j}. \end{aligned}$$

(if $\lambda_i \neq \lambda_j$, otherwise a similar computation produces $f'(\lambda_i)$.)

This completes the proof.

Condition number bound

If X is diagonalizable, we can replace the Schur form with an eigendecomposition, and obtain a bound

Theorem

Let $X = V\Lambda V^{-1}$ be diagonalizable. Then, for the Frobenius norm,

$$\kappa_{abs}(f, X) = \|\hat{L}_{f,X}\| \leq \kappa_2(V)^2 \max_{i,j} |f[\lambda_i, \lambda_j]|.$$

(And then as usual $\kappa_{rel}(f, X) = \kappa_{abs}(f, X)\frac{\|X\|}{\|f(X)\|}$.)

This bound shows two 'causes' of ill-conditioning:

- $|f[\lambda_i, \lambda_j]|$ is large, or
- $\kappa_2(V)$ is large (i.e., X very non-normal).

Example $f(x) = \sqrt{x}$ (principal square root): for which choices of $\Lambda(X)$ are the incremental rations $|f[\lambda_i, \lambda_j]|$ large?