Methods for general matrix functions

We now explore methods for matrix functions in general (not restricting to specific choices of f). [Higham book, Ch. 4]

Simple strategy: diagonalize $A = V\Lambda V^{-1}$, then compute

$$f(A) = Vf(\Lambda)V^{-1} = V \begin{bmatrix} f(\lambda_1) & & \\ & \ddots & \\ & & f(\lambda_m) \end{bmatrix} V^{-1}.$$

Works fine if A is symmetric/Hermitian/normal (and Q orthogonal). Otherwise, errors on $f(\lambda_i)$ (or in the diagonalization itself) are amplified by a factor $\kappa(V)$ — possibly much higher than the conditioning of the problem.

Example: sqrt of $\left[\begin{smallmatrix} 3 & -1 \\ 1 & 1 \end{smallmatrix}\right]$.

Alternative: do 'matrix algebra' directly, e.g., evaluate polynomials in matrix arguments.

Polynomial evaluation

How to evaluate polynomials in a matrix argument?

Unlike scalar polynomials, Horner method (i.e., $(\ldots((p_d A + p_{d-1})A + p_{d-2})A + \ldots)$ for matrix arguments is no better than 'direct' evaluation (build powers of A incrementally and sum them).

Even better: divide the terms into 'chunks' of size $\sqrt{d}$, e.g.,

$$(p_8 A^2 + p_7 A + p_6)(A^3)^2 + (p_5 A^2 + p_4 A + p_3)A^3 + (p_2 A^2 + p_1 A_1 + p_0).$$

(Paterson-Stockmayer method. — requires more storage though.)

Input:

$p_0, p_1, \dots, p_d$ $\quad A$

$$\left(\dots \left(\left(\left(p_d A + p_{d-1} I \right) A + p_{d-2} I \right) A \dots \right) \right)$$

1)

$F = p_0 \cdot I$

$B = I$

for $i = 1 : d$

$\quad B = B * A$

$\quad F = F + p_i * B$

end

↑

d molt. matrice-matrice

d add.

d matrice-scalare

$\sim d m^3$

2)

$F = p_d \cdot I + p_{d-1} I$

for $i = d : -1 : 1$

$\quad F = F * A$

$\quad F = F + p_{d-1} I$

end

↑

$\sim d$ moltiplicazioni

$\sim d$ addizioni

$\sim d m^3$

(Se ho ad es. una serie di Taylor troncata, meglio il metodo di sinistra perché non devo scegliere d all'inizio)

Paterson - Stockmeyer:

1) Calcola A^2 e A^3

$$\left(p_6 A^3 + p_5 A^2 + p_4 A\right) * A^3 + p_3 A^3 + p_2 A^2 + p_1 A + p_0 I$$

costa 3 moltiplicazioni: una per A^2, una per A^3, una in

"Spezzo" il polinomio in pezzi di lunghezza $\sqrt{d}$.

(quanti prodotti servono per grado 13? A: 6)

Padé approximations

Variant: Padé approximations, i.e., rational approximations.

Padé approximant (at $x = 0$)

For almost every f analytic at 0 and for every choice of degrees $\deg p, \deg q$, one can find a rational function $\frac{p(x)}{q(x)}$ such that

$$f(x) - \frac{p(x)}{q(x)} = \mathcal{O}(x^{\deg p + \deg q + 1}).$$

i.e., "matches first $\deg p + \deg q$ terms of the MacLaurin series". (Count degrees of freedom to get a hint of why it works.)

For many functions, they have better approximation properties than Taylor series.

We will examine them for specific functions, e.g. the square root.

$\dfrac{ax^2+bx+c}{dx^3+ex^2+fx+g}$ posso "raggiungere" i primi termini dello sviluppo di Taylor di una data funzione, scegliendo bene i parametri, ad es.

$$\frac{p(x)}{q(x)} = 1+x+\frac{1}{2}x^2+\frac{1}{3}x^3+ \dots + x^{\deg p+\deg q} + O\left(x^{\deg p+\deg q+1}\right)$$

$$= e^x + O\left(x^{\deg p + \deg q + 1}\right)$$

$\deg p+1$ coefficienti al numeratore

$\deg q+1$ al denominatore

-1 perché $\dfrac{p(x)}{q(x)} = \dfrac{\alpha p(x)}{\alpha q(x)}$.

In $q(x)\cdot\left(f(0)+f'(0)x+\frac{f''(0)}{2}x^2+\dots+\frac{f^{(k)}(0)}{k!}x^k\right)-p(x)$

potete scegliere i coeff. di p e q in modo da annullare i primi $\deg p+\deg q+1$ coefficienti del polinomio in x risultante ~~no~~ sono condizioni lineari nei coefficienti di $p(x)$ e $q(x)$.

A patto che la matrice $(\deg p+\deg q+1)\times(\deg p+\deg p+1)$ associata sia nonsingolare, e che $q(0)\neq 0$, riesco a trovare questi coefficienti.

Matrix approximants

Good approximation of a scalar function is not good enough: even if $|f(x) - p(x)| < \varepsilon$ for each x, this only implies

$\|V\| \cdot \|V^{-1}\|$

$$\|f(A) - p(A)\| = \|V(f(\Lambda) - p(\Lambda))V^{-1}\| \leq \kappa(V)\varepsilon.$$

One needs to study approximation properties directly "at the matrix level".

$$\left\| \begin{bmatrix} f(\lambda_1) - p(\lambda_1) & & \\ & \ddots & \\ & & f(\lambda_n) - p(\lambda_n) \end{bmatrix} \right\| \leq \varepsilon$$

Convergence of Taylor series

Theorem [Higham book Thm. 4.7]

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

$$\sum_{k=0}^{\infty} a_k (A-\alpha I)^k = \lim_{d\to\infty} \sum_{k=0}^{d} a_k (A-\alpha I)^k = f(A)$$

for each A whose eigenvalues satisfy $|\lambda_i - \alpha| < r$.

Proof (sketch):

- It is enough to work on Jordan blocks.
- If $p_d(x)$ is the polynomial obtained by truncating the series to degre-d, then $p_d(\lambda I + N) = \sum_{k=0}^{d} p_d^{(k)}(\lambda) N^k$.
- $p_d^{(k)}$ is the truncated Taylor series of $f^{(k)}$, which has the same radius of convergence as that of f. So $p_d^{(k)}(\lambda) \to f^{(k)}(\lambda)$.
- The sum has at most *size*(N) terms (all zero afterwards).

$A = V \cdot J \cdot V^{-1}$ $\quad J = \text{blkdiag}(J_1, J_2, \dots J_k)$

$$p(A) = V \,\text{blkdiag}\begin{pmatrix} p(J_1) & & \\ & \ddots & \\ & & p(J_k) \end{pmatrix} V^{-1}$$

Mi basta dire che $p_d(J_i) \to f(J_i)$, dove p_d è il polinomio di Taylor troncato al termine d.

Prendiamo un blocco di Jordan generico $J_i = \lambda_i I + N$ di dimensione $n_i \times n_i$.

$$p_d(J_i) = \sum_{k=0}^{d} a_k (\lambda_i I + N - \alpha I)^k = \sum_{k=0}^{d} b_k (\cancel{\lambda_i I} + N - \cancel{\lambda_i I})^k$$

$$= \sum_{k=0}^{d} b_k N^k = \sum_{k=0}^{n_i} b_k N^k \quad (*)$$

$$b_k = \frac{p_d^{(k)}(\lambda_i)}{k!}$$

La serie di Taylor di $f^{(k)}$ è data da

$$f^{(k)}(x) = f^{(k)}(\alpha) + f^{(k+1)}(\alpha)(x-\alpha) + \frac{f^{(k+2)}(\alpha)}{2!}(x-\alpha)^2 + \dots$$

ha raggio di convergenza uguale alla serie di Taylor di f, cioè r

$$b_k \to \frac{f^{(k)}(\lambda_i)}{k!}$$

$$P_d(J_i) = \sum_{k=0}^{n_i} b_k N^k \longrightarrow \sum_{k=0}^{n_i} \frac{f^{(k)}(\lambda_i)}{k!} N^k = f(J_i).$$

$f(x) = \frac{p(x)}{q(x)}$ con p, q polinomi, allora

$$f(A) = \underbrace{q(A)^{-1} p(A)} = \underbrace{p(A) q(A)^{-1}}$$

$$q(A) f(A) = (q \cdot f) A = p(A)$$

Parlett recurrence

Can one compute matrix functions using the Schur form of A?

triangolare perché è un polinomio in A

Example

$$A = \begin{bmatrix} t_{11} & t_{12} \\ 0 & t_{22} \end{bmatrix}, \quad f(A) = \begin{bmatrix} s_{11} & s_{12} \\ 0 & s_{22} \end{bmatrix}. = p(A) = \begin{bmatrix} p(t_{11}) & * \\ 0 & p(t_{22}) \end{bmatrix}$$

Clearly, $s_{11} = f(t_{11})$, $s_{22} = f(t_{22})$.

Trick: expanding $Af(A) = f(A)A$, one gets an equation for s_{12}:

$$\boxed{t_{11}s_{12} + t_{12}s_{22} = s_{11}t_{12} + s_{12}t_{22}} \implies s_{12} = t_{12}\boxed{\frac{s_{11} - s_{22}}{t_{11} - t_{22}}.}$$

(If $t_{11} = t_{22}$, the equation is not solvable and we already know that the finite difference becomes a derivative).

$$\begin{bmatrix} t_{11} & t_{12} \\ 0 & t_{22} \end{bmatrix}\begin{bmatrix} s_{11} & s_{12} \\ 0 & s_{22} \end{bmatrix} = \begin{bmatrix} s_{11} & s_{12} \\ 0 & s_{22} \end{bmatrix}\begin{bmatrix} t_{11} & t_{12} \\ 0 & t_{22} \end{bmatrix}$$

$$= t_{12}\frac{f(t_{11}) - f(t_{22})}{t_{11} - t_{22}}$$

Parlett recurrence — II

The same idea works for larger blocks (provided we compute things in the correct order): $A \cdot f(A) = f(A) \cdot A$

$$A = \begin{bmatrix} t_{11} & t_{12} & t_{13} \\ & t_{22} & t_{23} \\ & & t_{33} \end{bmatrix}, \quad f(A) = \begin{bmatrix} s_{11} & s_{12} & s_{13} \\ 0 & s_{22} & s_{23} \\ 0 & 0 & s_{33} \end{bmatrix},$$

$$\underline{t_{11}s_{13} + t_{12}s_{23} + t_{13}s_{33}} = \underline{s_{11}t_{13} + s_{12}t_{23} + s_{13}t_{33}}.$$

Very similar to the algorithm we used to solve Sylvester equations. In some sense, we are solving the (singular) Sylvester equation $\boxed{AX - XA = 0}$, after setting specific elements on its diagonal.

The same idea works blockwise — the quotients become Sylvester equations.

$$\begin{bmatrix} a_{11} & a_{12} \\ \boxed{0 \quad a_{22}} \end{bmatrix} \begin{bmatrix} \boxed{X_{11}} & X_{12} \\ \boxed{X_{21}} & X_{22} \end{bmatrix} + \begin{bmatrix} X_{11} & X_{12} \\ \boxed{X_{21} \quad X_{22}} \end{bmatrix} \begin{bmatrix} \boxed{b_{11}} & b_{12} \\ \boxed{0} & b_{22} \end{bmatrix} = \begin{bmatrix} C_{11} & C_{12} \\ \boxed{C_{21}} & C_{22} \end{bmatrix}$$

$$a_{22}X_{21} + X_{21}b_{11} = C_{21}$$

Se gli autovalori di A sono tutti distinti, riesco a calcolare iterativamente tutti gli S_{ij} fuori dalla diagonale

---

$$\begin{bmatrix} T_{11} & T_{12} \\ 0 & T_{22} \end{bmatrix} \begin{bmatrix} S_{11} & S_{12} \\ 0 & S_{22} \end{bmatrix} = \begin{bmatrix} S_{11} & S_{12} \\ 0 & S_{22} \end{bmatrix} \begin{bmatrix} T_{11} & T_{12} \\ 0 & T_{22} \end{bmatrix}$$

con blocchi della stessa dimensione, quadrati sulla diagonale

$T_{11}S_{12} + T_{12}S_{22} = S_{11}T_{12} + S_{12}T_{22}$ Se so già S_{11}, S_{22}, è un'eq. di Sylvester: $T_{11}S_{12} - S_{12}T_{22} = S_{11}T_{12} - T_{12}S_{22}$ $\leftarrow$

Ricorrenze: $\forall i, j$

$$t_{ii} S_{ij} + t_{i,i+1} S_{i+1,j} + \dots + t_{ij} S_{jj} = S_{ii} t_{ij} + S_{i,i+1} t_{i+1,j} + \dots + S_{ij} t_{jj}$$

se conosciamo S_{kj} per ogni $i < k \leq j$

e S_{ik} per ogni $i \leq k < j$,

possiamo risolvere come un'eq. di Sylvester per S_{ij}

$$t_{ii} S_{ij} - S_{ij} t_{jj} = \text{(altri termini)}$$

Risolubile se t_{ii}, t_{jj} non hanno autoval. in comune

Parlett recurrence — III

Algorithm (Schur–Parlett method)

1. Compute Schur form $A = QTQ^*$;
2. Partition T into blocks with 'well-separated eigenvalues';

2.5 Reorder (T, Q)

3. Compute $f(T_{ii})$ (e.g., with Taylor series in the centroid of its eigenvalues);
4. Use recurrences to compute off-diagonal blocks of $f(T)$;
5. Return $f(A) = Qf(T)Q^*$.

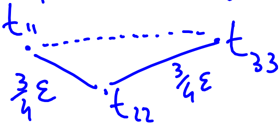
Tries to get 'best of both worlds': uses Taylor expansion when the eigenvalues are close, recurrences when they are distant.

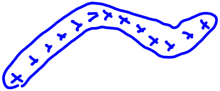
divisioni $t_{ii} - t_{jj}$ con t_{ii}, t_{jj} in blocchi diversi

Parto da $(t_{11}, t_{22}, \dots t_{nn})$; se $|t_{ii}-t_{jj}| \lesssim \varepsilon \approx 0.1$,

li metto nello stesso sottoinsieme.

Questo potrebbe raggruppare autovalori separati da più di ε, ma non ce ne preoccupiamo.

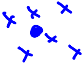




Però se t_{ii} e t_{jj} sono in insiemi diversi, sono ben separati. Riordino:

$$\begin{bmatrix} T_{11} & \text{Syl} & \text{Syl} \\ 0 & T_{22} & \text{Syl} \\ 0 & 0 & T_{33} \end{bmatrix}$$

in modo che autoval. nello stesso blocco siano consecutivi.



Parlett recurrence and block diagonalization

The Parlett recurrence is 'almost the same thing' as block diagonalization. Consider the case of 2 blocks for simplicity. T can be block-diagonalized via

$$W^{-1}TW = \begin{bmatrix} I & -X \\ 0 & I \end{bmatrix} \begin{bmatrix} T_{11} & T_{12} \\ 0 & T_{22} \end{bmatrix} \begin{bmatrix} I & X \\ 0 & I \end{bmatrix} = \begin{bmatrix} T_{11} & \\ & T_{22} \end{bmatrix}$$

where X solves $T_{11}X - XT_{22} + T_{12} = 0$ (Sylvester equation). Then

$$f(T) = W \begin{bmatrix} f(T_{11}) & \\ & f(T_{22}) \end{bmatrix} W^{-1} = \begin{bmatrix} f(T_{11}) & Xf(T_{22}) - f(T_{11})X \\ & f(T_{22}) \end{bmatrix}.$$

(Note indeed that $S = Xf(T_{22}) - f(T_{11})X$ solves the Sylvester equation appearing in the Parlett recurrence.)

So both methods solve a Sylvester equation with operator $Z \mapsto T_{11}Z - ZT_{22}$.

$$\underbrace{\begin{bmatrix} 1 & -X \\ 0 & 1 \end{bmatrix}}_{W^{-1}} \begin{bmatrix} T_{11} & T_{12} \\ 0 & T_{22} \end{bmatrix} \underbrace{\begin{bmatrix} 1 & X \\ 0 & 1 \end{bmatrix}}_{W} = \begin{bmatrix} T_{11} & 0 \\ 0 & T_{22} \end{bmatrix}$$

$$= \begin{bmatrix} T_{11} & T_{12} + T_{11}X - XT_{22} \\ 0 & T_{22} \end{bmatrix} \overset{!}{=} \begin{bmatrix} T_{11} & 0 \\ 0 & T_{22} \end{bmatrix}$$

so X solve $T_{11}X - XT_{22} + T_{12} = 0$

$$f\left(\begin{bmatrix} T_{11} & T_{12} \\ 0 & T_{22} \end{bmatrix}\right) = W f\left(W^{-1}\begin{bmatrix} T_{11} & T_{12} \\ 0 & T_{22} \end{bmatrix} W\right) W^{-1}$$

$$= \begin{bmatrix} 1 & X \\ 0 & 1 \end{bmatrix} f\left(\begin{bmatrix} T_{11} & 0 \\ 0 & T_{22} \end{bmatrix}\right)\begin{bmatrix} 1 & -X \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & X \\ 0 & 1 \end{bmatrix}\begin{bmatrix} f(T_{11}) & 0 \\ 0 & f(T_{22}) \end{bmatrix}\begin{bmatrix} 1 & -X \\ 0 & 1 \end{bmatrix} =$$

$$= \begin{bmatrix} 1 & X \\ 0 & 1 \end{bmatrix} \begin{bmatrix} f(T_{11}) & 0 \\ 0 & f(T_{22}) \end{bmatrix} \begin{bmatrix} 1 & -X \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} f(T_{11}) & \underline{Xf(T_{22}) - f(T_{11})X} \\ 0 & f(T_{22}) \end{bmatrix}$$

2 modi di trovare il blocco (1,2):

A. (Block-diag): Risolvi $\underline{T_{11}X - XT_{22}} + T_{12} = 0$, prendi $Xf(T_{22}) - f(T_{11})X$

B. (Schur-Parlett): Risolvi $\underline{T_{11}Y - YT_{22}} = f(T_{11})T_{12} - T_{12}f(T_{22})$

Matrice associata sempre $I \otimes T_{11} - T_{22}^T \otimes I$.

Teo: Questi metodi producono lo stesso risultato: basta dire che se X risolve $T_{11}X - XT_{22} + T_{12} = 0$, allora $Y := Xf(T_{22}) - f(T_{11})X$ risolve $T_{11}Y - YT_{22} = f(T_{11})T_{12} - T_{12}f(T_{22})$

$$T_{11}Y - YT_{22} = T_{11}\left(Xf(T_{22}) - f(T_{11})X\right) - \left(Xf(T_{22}) - f(T_{11})X\right)T_{22}$$

$$= T_{11}Xf(T_{22}) - \underbrace{T_{11}f(T_{11})}_{\text{commutano}}X - X\underbrace{f(T_{22})T_{22}}_{\text{commutano}} + f(T_{11})XT_{22} =$$

$$= \left(T_{11}X - XT_{22}\right)f(T_{22}) - f(T_{11})\left(T_{11}X - XT_{22}\right) =$$

$$= -T_{12}f(T_{22}) + f(T_{11})T_{12}$$