

# Come sfruttare la struttura di rango in un problema razionale agli autovalori

Gian Maria Negri Porzio

# Piano della presentazione

Introduzione

Cenni preliminari

Da REP a LEP

Applicazioni e risultati numerici

# Il problema degli autovalori e le sue generalizzazioni

## Definizione

Data una matrice  $A \in \mathbb{R}^{n \times n}$  e  $I$  la matrice identità chiamiamo **problema agli autovalori** la risoluzione dell'equazione  $\det(\lambda I - A) = 0$

Questo concetto è stato velocemente generalizzato:

- ▶ *Linear Eigenvalues Problem*:  $\det(\lambda A_1 + A_0) = 0$ .
- ▶ *Polynomial Eigenvalues Problem*:  $\det(\sum_{i=0}^n \lambda^i A_i) = 0$ .
- ▶ *Rational Eigenvalues Problem*:  $\det(P(\lambda) - \sum_{i=1}^k \frac{s_i(\lambda)}{q_i(\lambda)} A_i) = 0$ .
- ▶ *Nonlinear Eigenvalues Problem*:  $\det(f(\lambda)) = 0$ .

# Il problema degli autovalori e le sue generalizzazioni

## Definizione

Data una matrice  $A \in \mathbb{R}^{n \times n}$  e  $I$  la matrice identità chiamiamo **problema agli autovalori** la risoluzione dell'equazione  $\det(\lambda I - A) = 0$

Questo concetto è stato velocemente generalizzato:

- ▶ *Linear Eigenvalues Problem*:  $\det(\lambda A_1 + A_0) = 0$ .
- ▶ *Polynomial Eigenvalues Problem*:  $\det(\sum_{i=0}^n \lambda^i A_i) = 0$ .
- ▶ *Rational Eigenvalues Problem*:  $\det(P(\lambda) - \sum_{i=1}^k \frac{s_i(\lambda)}{q_i(\lambda)} A_i) = 0$ .
- ▶ *Nonlinear Eigenvalues Problem*:  $\det(f(\lambda)) = 0$ .

# Il problema degli autovalori e le sue generalizzazioni

## Definizione

Data una matrice  $A \in \mathbb{R}^{n \times n}$  e  $I$  la matrice identità chiamiamo **problema agli autovalori** la risoluzione dell'equazione  $\det(\lambda I - A) = 0$

Questo concetto è stato velocemente generalizzato:

- ▶ *Linear Eigenvalues Problem*:  $\det(\lambda A_1 + A_0) = 0$ .
- ▶ *Polynomial Eigenvalues Problem*:  $\det(\sum_{i=0}^n \lambda^i A_i) = 0$ .
- ▶ *Rational Eigenvalues Problem*:  $\det(P(\lambda) - \sum_{i=1}^k \frac{s_i(\lambda)}{q_i(\lambda)} A_i) = 0$ .
- ▶ *Nonlinear Eigenvalues Problem*:  $\det(f(\lambda)) = 0$ .

# Il problema degli autovalori e le sue generalizzazioni

## Definizione

Data una matrice  $A \in \mathbb{R}^{n \times n}$  e  $I$  la matrice identità chiamiamo **problema agli autovalori** la risoluzione dell'equazione  $\det(\lambda I - A) = 0$

Questo concetto è stato velocemente generalizzato:

- ▶ *Linear Eigenvalues Problem*:  $\det(\lambda A_1 + A_0) = 0$ .
- ▶ *Polynomial Eigenvalues Problem*:  $\det(\sum_{i=0}^n \lambda^i A_i) = 0$ .
- ▶ *Rational Eigenvalues Problem*:  $\det(P(\lambda) - \sum_{i=1}^k \frac{s_i(\lambda)}{q_i(\lambda)} A_i) = 0$ .
- ▶ *Nonlinear Eigenvalues Problem*:  $\det(f(\lambda)) = 0$ .

## Ipotesi di lavoro

Da questo momento lavoreremo su una matrice di funzioni razionali  $R(\lambda)$  di dimensione  $n \times n$  :

$$R(\lambda) = P(\lambda) - \sum_{i=1}^k \frac{s_i(\lambda)}{q_i(\lambda)} E_i$$

tale che:

- ▶  $\det(R(\lambda)) \neq 0$ .
- ▶  $P(\lambda) = \sum_{i=0}^d \lambda^i A_i$ ,  $A_d = I$ .
- ▶  $\text{rk } E_i = r_i$ .
- ▶  $s_i(\lambda)$ ,  $q_i(\lambda)$  sono coprimi e  $\deg(s_i) < \deg(q_i) = d_i$  .

## Un approccio intuitivo

Un approccio *brute-force* consiste nel moltiplicare  $R(\lambda)$  per il minimo comune multiplo di  $\{q_i(\lambda) \mid 1 \leq i \leq k\}$  e ricondurci a un PEP da linearizzare in un secondo momento.

Questo comporta due notevoli svantaggi:

- ▶ L'aumento delle dimensioni del problema.
- ▶ La possibile perdita della struttura di rango.

### Soluzione

Yangfeng Su e Zhaojun Bai hanno proposto una linearizzazione che sfrutti intensamente le proprietà di rango basso delle matrici  $E_i$ .

## Un approccio intuitivo

Un approccio *brute-force* consiste nel moltiplicare  $R(\lambda)$  per il minimo comune multiplo di  $\{q_i(\lambda) \mid 1 \leq i \leq k\}$  e ricondurci a un PEP da linearizzare in un secondo momento.

Questo comporta due notevoli svantaggi:

- ▶ L'aumento delle dimensioni del problema.
- ▶ La possibile perdita della struttura di rango.

### Soluzione

Yangfeng Su e Zhaojun Bai hanno proposto una linearizzazione che sfrutti intensamente le proprietà di rango basso delle matrici  $E_i$ .

## Un approccio intuitivo

Un approccio *brute-force* consiste nel moltiplicare  $R(\lambda)$  per il minimo comune multiplo di  $\{q_i(\lambda) \mid 1 \leq i \leq k\}$  e ricondurci a un PEP da linearizzare in un secondo momento.

Questo comporta due notevoli svantaggi:

- ▶ L'aumento delle dimensioni del problema.
- ▶ La possibile perdita della struttura di rango.

### Soluzione

Yangfeng Su e Zhaojun Bai hanno proposto una linearizzazione che sfrutti intensamente le proprietà di rango basso delle matrici  $E_i$ .

# Minimal realization

Prima di costruire la nostra linearizzazione avremo bisogno di due lemmi.

## Lemma

*Siano  $s(\lambda)$  e  $q(\lambda)$  due polinomi come da ipotesi precedenti. Allora esistono due vettori  $a, b \in \mathbb{R}^d$ , una matrice  $C \in \mathbb{R}^{d \times d}$  e una matrice  $D \in \mathbb{R}^{d \times d}$  diagonale invertibile tali che:*

$$\frac{s(\lambda)}{q(\lambda)} = a^T (C - \lambda D)^{-1} b$$

## Minimal realization

Prima di costruire la nostra linearizzazione avremo bisogno di due lemmi.

### Lemma

*Siano  $s(\lambda)$  e  $q(\lambda)$  due polinomi come da ipotesi precedenti. Allora esistono due vettori  $a, b \in \mathbb{R}^d$ , una matrice  $C \in \mathbb{R}^{d \times d}$  e una matrice  $D \in \mathbb{R}^{d \times d}$  diagonale invertibile tali che:*

$$\frac{s(\lambda)}{q(\lambda)} = a^T (C - \lambda D)^{-1} b$$

## Minimal realization

### Dimostrazione.

Scriviamo  $s(\lambda) = \sum_{j=0}^{d-1} a_j \lambda^j$  e  $q(\lambda) = \sum_{j=0}^d b_j \lambda^j$ . Allora l'uguaglianza viene verificata semplicemente da:

$$\frac{s(\lambda)}{q(\lambda)} = [a_{d-1} \quad a_{d-2} \quad \cdots \quad a_0] \begin{bmatrix} \lambda + \frac{b_{d-1}}{b_d} & \frac{b_{d-2}}{b_d} & \cdots & \frac{b_0}{b_d} \\ -1 & \lambda & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & -1 & \lambda \end{bmatrix}^{-1} \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

da cui ricaviamo i valori di  $C$ ,  $D$ ,  $a$ ,  $b$ . □

## Decomposizione rank revealing

Assumiamo infine di poter decomporre le matrici  $E_i \in \mathbb{R}^{n \times n}$  di rango  $r_i$  nella forma:

$$E_i = L_i U_i^T$$

con  $L_i, U_i$  matrici di dimensione  $n \times r_i$ .

### Osservazione

Questo sarà il **passaggio chiave** della nostra costruzione

## Decomposizione rank revealing

Assumiamo infine di poter decomporre le matrici  $E_i \in \mathbb{R}^{n \times n}$  di rango  $r_i$  nella forma:

$$E_i = L_i U_i^T$$

con  $L_i, U_i$  matrici di dimensione  $n \times r_i$ .

### Osservazione

Questo sarà il **passaggio chiave** della nostra costruzione

## Costruzione delle matrici pencil

Ricordiamo la nostra situazione. Abbiamo un REP:

$$R(\lambda) = P(\lambda) - \sum_{i=1}^k \frac{s_i(\lambda)}{q_i(\lambda)} E_i$$

e vogliamo ricondurci ad un LEP della forma:

$$\mathcal{A} - \lambda \mathcal{B}$$

Iniziamo considerando solamente  $\sum_{i=1}^k \frac{s_i(\lambda)}{q_i(\lambda)} E_i$ .

## Costruzione delle matrici pencil

Ricordiamo la nostra situazione. Abbiamo un REP:

$$R(\lambda) = P(\lambda) - \sum_{i=1}^k \frac{s_i(\lambda)}{q_i(\lambda)} E_i$$

e vogliamo ricondurci ad un LEP della forma:

$$\mathcal{A} - \lambda \mathcal{B}$$

Iniziamo considerando solamente  $\sum_{i=1}^k \frac{s_i(\lambda)}{q_i(\lambda)} E_i$ .

## Costruzione delle matrici pencil

Per quanto visto nelle diapositive precedenti:

$$\begin{aligned}
 \sum_{i=1}^k \frac{s_i(\lambda)}{q_i(\lambda)} E_i &= \sum_{i=1}^k \left[ a_i^T (C_i - \lambda D_i)^{-1} b_i \right] L_i U_i^T \\
 &= \sum_{i=1}^k L_i \left[ a_i^T (C_i - \lambda D_i)^{-1} b_i \cdot l_{r_i} \right] U_i^T \\
 &= \sum_{i=1}^k L_i (l_{r_i} \otimes a_i)^T (l_{r_i} \otimes C_i - l_{r_i} \otimes D_i)^{-1} (l_{r_i} \otimes b_i) U_i^T
 \end{aligned}$$

## Costruzione delle matrici pencil

Definendo:

- ▶  $C = \text{diag}(l_{r_1} \otimes C_1, l_{r_2} \otimes C_2, \dots, l_{r_k} \otimes C_k),$
- ▶  $D = \text{diag}(l_{r_1} \otimes D_1, l_{r_2} \otimes D_2, \dots, l_{r_k} \otimes D_k),$
- ▶  $L = [L_1(l_{r_1} \otimes a_1)^T, L_2(l_{r_2} \otimes a_2)^T, \dots, L_k(l_{r_k} \otimes a_k)^T],$
- ▶  $U = [U_1(l_{r_1} \otimes b_1)^T, U_2(l_{r_2} \otimes b_2)^T, \dots, U_k(l_{r_k} \otimes b_k)^T],$

Possiamo scrivere:

$$\begin{aligned} \sum_{i=1}^k \frac{s_i(\lambda)}{q_i(\lambda)} E_i &= \sum_{i=1}^k L_i(l_{r_i} \otimes a_i)^T (l_{r_i} \otimes C_i - \lambda D)^{-1} (l_{r_i} \otimes b_i) U_i^T \\ &= L(C - \lambda D)^{-1} U^T \end{aligned}$$

Si noti che come gli autovalori  $(C - \lambda D)$  siano i poli di  $R(\lambda)$



## Costruzione delle matrici pencil

Definendo:

- ▶  $C = \text{diag}(I_{r_1} \otimes C_1, I_{r_2} \otimes C_2, \dots, I_{r_k} \otimes C_k),$
- ▶  $D = \text{diag}(I_{r_1} \otimes D_1, I_{r_2} \otimes D_2, \dots, I_{r_k} \otimes D_k),$
- ▶  $L = [L_1(I_{r_1} \otimes a_1)^T, L_2(I_{r_2} \otimes a_2)^T, \dots, L_k(I_{r_k} \otimes a_k)^T],$
- ▶  $U = [U_1(I_{r_1} \otimes b_1)^T, U_2(I_{r_2} \otimes b_2)^T, \dots, U_k(I_{r_k} \otimes b_k)^T],$

Possiamo scrivere:

$$\begin{aligned} \sum_{i=1}^k \frac{s_i(\lambda)}{q_i(\lambda)} E_i &= \sum_{i=1}^k L_i(I_{r_i} \otimes a_i)^T (I_{r_i} \otimes C_i - \lambda D)^{-1} (I_{r_i} \otimes b_i) U_i^T \\ &= L(C - \lambda D)^{-1} U^T \end{aligned}$$

Si noti che come gli autovalori  $(C - \lambda D)$  siano i poli di  $R(\lambda)$



# Costruzione delle matrici pencil

## Caso lineare

Supponiamo che  $P(\lambda) = A - \lambda B$ . Il nostro problema diventa

$$[A - \lambda B - L(C - \lambda D)^{-1}U^T]x = 0;$$

che possiamo riscrivere nella forma

$$(A - \lambda B)z = 0$$

con

$$A = \begin{bmatrix} A & L \\ U^T & C \end{bmatrix}, \quad B = \begin{bmatrix} B & \\ & D \end{bmatrix}$$

e  $z$  vettore opportuno

# Costruzione delle matrici pencil

## Caso lineare

Supponiamo che  $P(\lambda) = A - \lambda B$ . Il nostro problema diventa

$$[A - \lambda B - L(C - \lambda D)^{-1}U^T]x = 0;$$

che possiamo riscrivere nella forma

$$(A - \lambda B)z = 0$$

con

$$A = \begin{bmatrix} A & L \\ U^T & C \end{bmatrix}, \quad B = \begin{bmatrix} B & \\ & D \end{bmatrix}$$

e  $z$  vettore opportuno

# Costruzione delle matrici pencil

Caso generale

Se  $P(\lambda) = \sum_{i=0}^d A_i \lambda^i$  un ragionamento analogo ci conduce a

$$A = \left[ \begin{array}{cccc|c} A_{d-1} & A_{d-2} & \cdots & A_0 & L \\ -I & 0 & \cdots & 0 & \\ & & \ddots & \vdots & \\ & & & -I & 0 \\ \hline & & & U^T & C \end{array} \right]$$

$$B = - \left[ \begin{array}{ccc|c} A_d & & & \\ & I & & \\ & & \ddots & \\ & & & I \\ \hline & & & -D \end{array} \right].$$

# Considerazioni

Vantaggi:

- ▶ Nessun autovalore infinito.
- ▶ Dimensione delle matrici pencil.
- ▶ Versatilità: e.g. simmetria della minimal realization:

$$a_i^T (C_i - \lambda D)^{-1} b_i$$

# Correttezza dell'algoritmo

## Problema

Questa linearizzazione è corretta?

# Correttezza dell'algoritmo

## Teorema

Consideriamo il REP  $R(\lambda) = P(\lambda) - \sum_{i=0}^k \frac{s_i(\lambda)}{q_i(\lambda)} E_i$  e il LEP corrispondente  $\mathcal{A} - \lambda\mathcal{B}$ . Allora:

- 1  $\lambda$  autovalore di REP  $\Rightarrow \lambda$  autovalore di LEP.
- 2 Siano  $\lambda$  un autovalore di LEP ma NON un polo di  $R(\lambda)$  e

$$z = [z_1^T \quad z_2^T \quad \cdots \quad z_d^T \quad y^T]$$

il corrispondente autovettore. Allora  $(\lambda, z_d)$  è una coppia autovalore/vettore per REP.

## Molla caricata

Consideriamo il seguente problema tratto dalla raccolta NLEVP:

$$R(\lambda) = A - \lambda B + E + \frac{\sigma}{\lambda - \sigma} E. \quad (1)$$

con  $A, B$  matrici  $n \times n$  simmetriche, tridiagonali e definite positive,  $E = e_n e_n^T$  e  $\sigma > 0$ .

L'approccio *brute force* conduce al problema quadratico:

$$Q(\lambda) = \lambda^2 B - \lambda(A + \sigma B + E) + \sigma A \quad (2)$$

## Molla caricata

Consideriamo il seguente problema tratto dalla raccolta NLEVP:

$$R(\lambda) = A - \lambda B + E + \frac{\sigma}{\lambda - \sigma} E. \quad (1)$$

con  $A, B$  matrici  $n \times n$  simmetriche, tridiagonali e definite positive,  $E = e_n e_n^T$  e  $\sigma > 0$ .

L'approccio *brute force* conduce al problema quadratico:

$$Q(\lambda) = \lambda^2 B - \lambda(A + \sigma B + E) + \sigma A \quad (2)$$

## Linearizzazioni

Con la nostra linearizzazione:

$$\begin{bmatrix} A + e_n^T e_n & -\sigma e_n \\ -e_n & -\sigma \end{bmatrix} - \lambda \begin{bmatrix} B & \\ & -1 \end{bmatrix} \quad (3)$$

Con l'approccio intuitivo:

$$\begin{bmatrix} -(A + \sigma B e_n^T e_n) & \sigma A \\ \sigma A & \end{bmatrix} - \lambda \begin{bmatrix} -B & \\ & \sigma A \end{bmatrix} \quad (4)$$

Si noti di quanto cala la dimensione della linearizzazione: passiamo da  $2n$  a  $n + 1$ !

## Linearizzazioni

Con la nostra linearizzazione:

$$\begin{bmatrix} A + e_n^T e_n & -\sigma e_n \\ -e_n & -\sigma \end{bmatrix} - \lambda \begin{bmatrix} B & \\ & -1 \end{bmatrix} \quad (3)$$

Con l'approccio intuitivo:

$$\begin{bmatrix} -(A + \sigma B e_n^T e_n) & \sigma A \\ \sigma A & \end{bmatrix} - \lambda \begin{bmatrix} -B & \\ & \sigma A \end{bmatrix} \quad (4)$$

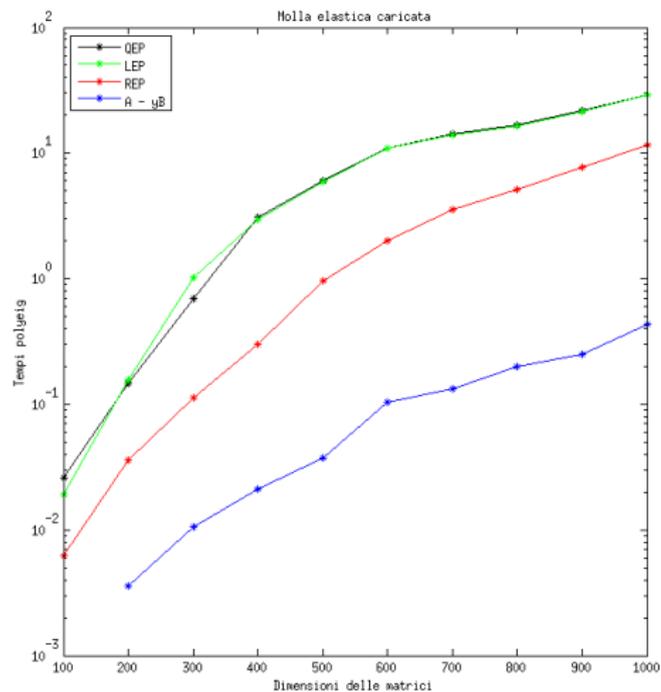
Si noti di quanto cala la dimensione della linearizzazione: passiamo da  $2n$  a  $n + 1$ !

## Prestazioni

Possiamo osservare l'aumento delle prestazioni per questa linearizzazione. Riportiamo i dati dei diversi approci al problema.

	$n = 100$	$n = 400$	$n = 700$	$n = 1000$
QEP (2)	0.0260	3.1033	14.0997	29.2851
LEP (4)	0.0192	2.9593	13.8810	29.0894
REP (3)	0.0062	0.3013	3.5440	11.6928
$A - \lambda B$	$\epsilon$	0.0211	0.1334	0.4366

# Prestazioni



## Un problema quadratico di rango basso

Sia

$$Q(\lambda) = \lambda^2 M + \lambda D + K$$

un problema quadratico  $n \times n$  con  $K$  matrice di rango  $r \ll n$ . Se siamo interessati agli autovalori non nulli possiamo trasformarlo in

$$R(\lambda) = \lambda M + D + \frac{1}{\lambda} K$$

e sfruttarne la struttura.

### Esempio

Il problema *railtrack* della raccolta NLEVP.

## Un problema quadratico di rango basso

Sia

$$Q(\lambda) = \lambda^2 M + \lambda D + K$$

un problema quadratico  $n \times n$  con  $K$  matrice di rango  $r \ll n$ . Se siamo interessati agli autovalori non nulli possiamo trasformarlo in

$$R(\lambda) = \lambda M + D + \frac{1}{\lambda} K$$

e sfruttarne la struttura.

### Esempio

Il problema *railtrack* della raccolta NLEVP.

## Un problema quadratico di rango basso

Sia

$$Q(\lambda) = \lambda^2 M + \lambda D + K$$

un problema quadratico  $n \times n$  con  $K$  matrice di rango  $r \ll n$ . Se siamo interessati agli autovalori non nulli possiamo trasformarlo in

$$R(\lambda) = \lambda M + D + \frac{1}{\lambda} K$$

e sfruttarne la struttura.

### Esempio

Il problema *railtrack* della raccolta NLEVP.

# Riassumendo

- 1 Abbiamo sviluppato una linearizzazione che sfrutta la struttura di rango.
- 2 Ne abbiamo mostrato i vantaggi computazionali.
- 3 Abbiamo elencato alcune applicazioni pratiche.