

Localizzazione di zeri di funzioni analitiche mediante integrali di contorno

Igor Simunec

Relatore: Dario Andrea Bini

Dipartimento di Matematica
Università di Pisa

13 luglio 2018

- 1 Presentazione del problema
- 2 Risultati teorici
- 3 Algoritmo
- 4 Varianti
- 5 Interpolazione razionale

- 1 Presentazione del problema
- 2 Risultati teorici
- 3 Algoritmo
- 4 Varianti
- 5 Interpolazione razionale

Presentazione del problema

Consideriamo una funzione analitica $f : U \rightarrow \mathbb{C}$ definita su un aperto semplicemente connesso $U \subset \mathbb{C}$.

Il nostro scopo è determinare gli zeri di f in una regione $D \subset U$ delimitata da una curva semplice γ .

Presentazione del problema

Consideriamo una funzione analitica $f : U \rightarrow \mathbb{C}$ definita su un aperto semplicemente connesso $U \subset \mathbb{C}$.

Il nostro scopo è determinare gli zeri di f in una regione $D \subset U$ delimitata da una curva semplice γ .

Notazione:

- f ha n zeri distinti in D , indicati con z_1, \dots, z_n ;
- lo zero z_k ha molteplicità ν_k ;
- il numero totale di zeri è $N := \sum_{k=1}^n \nu_k$.

Presentazione del problema

Consideriamo una funzione analitica $f : U \rightarrow \mathbb{C}$ definita su un aperto semplicemente connesso $U \subset \mathbb{C}$.

Il nostro scopo è determinare gli zeri di f in una regione $D \subset U$ delimitata da una curva semplice γ .

Notazione:

- f ha n zeri distinti in D , indicati con z_1, \dots, z_n ;
- lo zero z_k ha molteplicità ν_k ;
- il numero totale di zeri è $N := \sum_{k=1}^n \nu_k$.

In generale n non è noto, mentre N può essere calcolato mediante il teorema dei residui:

$$N = \frac{1}{2\pi i} \int_{\gamma} \frac{f'(z)}{f(z)} dz.$$

- 1 Presentazione del problema
- 2 Risultati teorici**
- 3 Algoritmo
- 4 Varianti
- 5 Interpolazione razionale

Definizione

Definiamo il prodotto scalare $\langle \cdot, \cdot \rangle : \mathbb{C}[z] \times \mathbb{C}[z] \rightarrow \mathbb{C}$ mediante un integrale su γ . Per ϕ, ψ polinomi sia

$$\langle \phi, \psi \rangle := \frac{1}{2\pi i} \int_{\gamma} \phi(z) \psi(z) \frac{f'(z)}{f(z)} dz = \sum_{k=1}^n \nu_k \phi(z_k) \psi(z_k).$$

Definizione

Definiamo il prodotto scalare $\langle \cdot, \cdot \rangle : \mathbb{C}[z] \times \mathbb{C}[z] \rightarrow \mathbb{C}$ mediante un integrale su γ . Per ϕ, ψ polinomi sia

$$\langle \phi, \psi \rangle := \frac{1}{2\pi i} \int_{\gamma} \phi(z)\psi(z) \frac{f'(z)}{f(z)} dz = \sum_{k=1}^n \nu_k \phi(z_k)\psi(z_k).$$

Introduciamo i momenti $s_p := \langle 1, z^p \rangle = \sum_{k=1}^n \nu_k z_k^p$ e le matrici di Hankel

$$H_t := [s_{p+q}]_{p,q=0}^{t-1} = \begin{bmatrix} s_0 & \dots & s_{t-1} \\ \vdots & \ddots & \vdots \\ s_{t-1} & \dots & s_{2t-2} \end{bmatrix}.$$

Definizione (FOP)

Un polinomio monico ϕ_t di grado $t \geq 0$ si dice *polinomio ortogonale formale* (FOP) se soddisfa la condizione

$$\langle z^k, \phi_t(z) \rangle = 0, \quad k = 0, 1, \dots, t - 1. \quad (1)$$

Se ϕ_t esiste ed è unico, è detto un FOP regolare e t un indice regolare.

Definizione (FOP)

Un polinomio monico ϕ_t di grado $t \geq 0$ si dice *polinomio ortogonale formale* (FOP) se soddisfa la condizione

$$\langle z^k, \phi_t(z) \rangle = 0, \quad k = 0, 1, \dots, t - 1. \quad (1)$$

Se ϕ_t esiste ed è unico, è detto un FOP regolare e t un indice regolare.

Osservazione

L'indice t è regolare se e solo se la matrice H_t è non singolare.

Lemma

Per ogni $p \geq 0$ vale che $\text{rank } H_{n+p} = n$.

Lemma

Per ogni $p \geq 0$ vale che $\text{rank } H_{n+p} = n$.

- Il polinomio $\phi_n(z) := \prod_{k=1}^n (z - z_k)$ è ortogonale a tutti i polinomi, quindi è il FOP regolare di indice n .
- Se vale $s_0 = N \neq 0$, possiamo definire $\mu := s_1/s_0$ e abbiamo che $\phi_1(z) = z - \mu$ è il FOP regolare di indice 1.

Lemma

Per ogni $p \geq 0$ vale che $\text{rank } H_{n+p} = n$.

- Il polinomio $\phi_n(z) := \prod_{k=1}^n (z - z_k)$ è ortogonale a tutti i polinomi, quindi è il FOP regolare di indice n .
- Se vale $s_0 = N \neq 0$, possiamo definire $\mu := s_1/s_0$ e abbiamo che $\phi_1(z) = z - \mu$ è il FOP regolare di indice 1.

Definizione (*Inner polynomials*)

Sia t un indice non regolare. Un *inner polynomial* di indice t è un polinomio $\phi_t := \phi_r \psi_{t,r}$ tale che

- r è il più grande indice regolare minore di t ;
- $\psi_{t,r}$ è un polinomio monico di grado $t - r$, per esempio z^{t-r} .

Usando la successione di FOP $\{\phi_t\}_{t \in \mathbb{N}}$, possiamo definire le matrici

$$G_t := [\langle \phi_r, \phi_s \rangle]_{r,s=0}^{t-1} \quad \text{e} \quad G_t^{(1)} := [\langle \phi_1 \phi_r, \phi_s \rangle]_{r,s=0}^{t-1}.$$

Usando la successione di FOP $\{\phi_t\}_{t \in \mathbb{N}}$, possiamo definire le matrici

$$G_t := [\langle \phi_r, \phi_s \rangle]_{r,s=0}^{t-1} \quad \text{e} \quad G_t^{(1)} := [\langle \phi_1 \phi_r, \phi_s \rangle]_{r,s=0}^{t-1}.$$

Il risultato principale su cui si basa l'algoritmo che presenteremo è il seguente:

Teorema

Sia $t \geq 1$ un indice regolare e siano $z_{t,1}, \dots, z_{t,t}$ gli zeri del FOP ϕ_t . Allora gli autovalori del pencil $G_t^{(1)} - \lambda G_t$ sono dati da $z_{t,1} - \mu, \dots, z_{t,t} - \mu$, dove $\mu = s_1/s_0$.

Usando la successione di FOP $\{\phi_t\}_{t \in \mathbb{N}}$, possiamo definire le matrici

$$G_t := [\langle \phi_r, \phi_s \rangle]_{r,s=0}^{t-1} \quad \text{e} \quad G_t^{(1)} := [\langle \phi_1 \phi_r, \phi_s \rangle]_{r,s=0}^{t-1}.$$

Il risultato principale su cui si basa l'algoritmo che presenteremo è il seguente:

Teorema

Sia $t \geq 1$ un indice regolare e siano $z_{t,1}, \dots, z_{t,t}$ gli zeri del FOP ϕ_t . Allora gli autovalori del pencil $G_t^{(1)} - \lambda G_t$ sono dati da $z_{t,1} - \mu, \dots, z_{t,t} - \mu$, dove $\mu = s_1/s_0$.

Corollario

Gli zeri di f coincidono con gli autovalori di $G_n^{(1)} - \lambda G_n$, traslati di μ .

- 1 Presentazione del problema
- 2 Risultati teorici
- 3 Algoritmo**
- 4 Varianti
- 5 Interpolazione razionale

Questo algoritmo è stato proposto da Kravanja, Sakurai e Van Barel (KSV) in [1].

- I prodotti scalari $\langle \phi_r, \phi_s \rangle$ vengono calcolati con metodi di integrazione numerica lungo γ . Se γ è una circonferenza, una scelta efficace è il metodo dei trapezi.

Questo algoritmo è stato proposto da Kravanja, Sakurai e Van Barel (KSV) in [1].

- I prodotti scalari $\langle \phi_r, \phi_s \rangle$ vengono calcolati con metodi di integrazione numerica lungo γ . Se γ è una circonferenza, una scelta efficace è il metodo dei trapezi.
- Se r è un indice regolare, possiamo determinare gli zeri del FOP ϕ_r avendo a disposizione le matrici $G_r^{(1)}$ e G_r , che contengono soltanto i prodotti scalari dei FOP di grado $\leq r - 1$.

Questo algoritmo è stato proposto da Kravanja, Sakurai e Van Barel (KSV) in [1].

- I prodotti scalari $\langle \phi_r, \phi_s \rangle$ vengono calcolati con metodi di integrazione numerica lungo γ . Se γ è una circonferenza, una scelta efficace è il metodo dei trapezi.
- Se r è un indice regolare, possiamo determinare gli zeri del FOP ϕ_r avendo a disposizione le matrici $G_r^{(1)}$ e G_r , che contengono soltanto i prodotti scalari dei FOP di grado $\leq r - 1$.
- Perciò i FOP ϕ_t , $t = 1, \dots, n$ possono essere calcolati in sequenza, risolvendo una successione di problemi generalizzati agli autovalori.

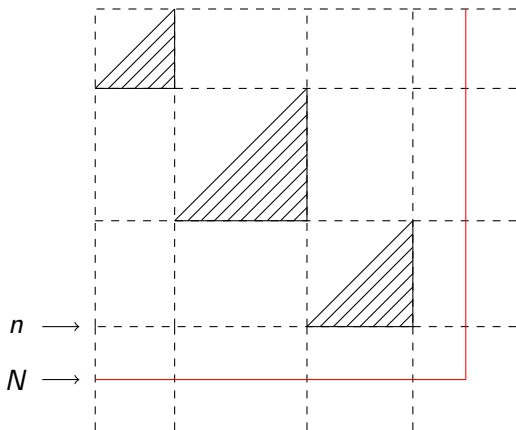
Questo algoritmo è stato proposto da Kravanja, Sakurai e Van Barel (KSV) in [1].

- I prodotti scalari $\langle \phi_r, \phi_s \rangle$ vengono calcolati con metodi di integrazione numerica lungo γ . Se γ è una circonferenza, una scelta efficace è il metodo dei trapezi.
- Se r è un indice regolare, possiamo determinare gli zeri del FOP ϕ_r avendo a disposizione le matrici $G_r^{(1)}$ e G_r , che contengono soltanto i prodotti scalari dei FOP di grado $\leq r - 1$.
- Perciò i FOP ϕ_t , $t = 1, \dots, n$ possono essere calcolati in sequenza, risolvendo una successione di problemi generalizzati agli autovalori.

Affinché l'algoritmo possa funzionare è necessario:

- 1 Saper decidere quando t è un indice regolare.
- 2 Saper riconoscere quando è stato raggiunto l'indice n .

Entrambe queste verifiche possono essere fatte osservando la struttura della matrice G_t :



Una volta calcolati gli zeri distinti z_1, \dots, z_n di f , le rispettive molteplicità si possono calcolare risolvendo il seguente sistema di Vandermonde:

$$\begin{bmatrix} 1 & 1 & \dots & 1 \\ z_1 & z_2 & \dots & z_n \\ \vdots & \vdots & & \vdots \\ z_1^{n-1} & z_2^{n-1} & \dots & z_n^{n-1} \end{bmatrix} \begin{bmatrix} \nu_1 \\ \nu_2 \\ \vdots \\ \nu_n \end{bmatrix} = \begin{bmatrix} s_0 \\ s_1 \\ \vdots \\ s_{n-1} \end{bmatrix} .$$

Esistono algoritmi che risolvono sistemi di questo tipo in tempo $\mathcal{O}(n^2)$, come quello di Gohberg e Koltracht [4].

- 1 Presentazione del problema
- 2 Risultati teorici
- 3 Algoritmo
- 4 Varianti**
- 5 Interpolazione razionale

Zeri in *clusters*

Supponiamo che gli zeri z_1, \dots, z_n di f si possano raggruppare in m *clusters*, con dimensione massima $\delta \ll 1$.

Supponiamo che gli zeri z_1, \dots, z_n di f si possano raggruppare in m *clusters*, con dimensione massima $\delta \ll 1$.

Allora vale il seguente risultato:

Teorema

L'indice m è regolare se $\delta \ll 1$. Inoltre per $\delta \rightarrow 0$ vale che:

- $\phi_m(c_j) = \mathcal{O}(\delta^2)$, dove c_1, \dots, c_m sono i centri dei *clusters*;
- $\langle z^p, \phi_m(z) \rangle = \mathcal{O}(\delta^2)$ per ogni $p \in \mathbb{N}$.

Supponiamo che gli zeri z_1, \dots, z_n di f si possano raggruppare in m *clusters*, con dimensione massima $\delta \ll 1$.

Allora vale il seguente risultato:

Teorema

L'indice m è regolare se $\delta \ll 1$. Inoltre per $\delta \rightarrow 0$ vale che:

- $\phi_m(c_j) = \mathcal{O}(\delta^2)$, dove c_1, \dots, c_m sono i centri dei *clusters*;
- $\langle z^p, \phi_m(z) \rangle = \mathcal{O}(\delta^2)$ per ogni $p \in \mathbb{N}$.

Quindi se applichiamo l'algoritmo KSV con una soglia ϵ_{stop} non troppo piccola, esso si arresta all'indice m e fornisce un'approssimazione dei centri c_1, \dots, c_m .

Supponiamo che gli zeri z_1, \dots, z_n di f si possano raggruppare in m *clusters*, con dimensione massima $\delta \ll 1$.

Allora vale il seguente risultato:

Teorema

L'indice m è regolare se $\delta \ll 1$. Inoltre per $\delta \rightarrow 0$ vale che:

- $\phi_m(c_j) = \mathcal{O}(\delta^2)$, dove c_1, \dots, c_m sono i centri dei *clusters*;
- $\langle z^p, \phi_m(z) \rangle = \mathcal{O}(\delta^2)$ per ogni $p \in \mathbb{N}$.

Quindi se applichiamo l'algoritmo KSV con una soglia ϵ_{stop} non troppo piccola, esso si arresta all'indice m e fornisce un'approssimazione dei centri c_1, \dots, c_m .

Per determinare con precisione gli zeri che compongono il *cluster* j -esimo, possiamo riapplicare l'algoritmo integrando su una piccola circonferenza γ_j centrata in c_j .

Alternativa *derivative-free*

Si può produrre una versione dell'algoritmo KSV che non richiede di valutare la derivata di f . Questo è utile nel caso in cui f' non è disponibile, o il suo calcolo è particolarmente costoso.

Alternativa *derivative-free*

Si può produrre una versione dell'algoritmo KSV che non richiede di valutare la derivata di f . Questo è utile nel caso in cui f' non è disponibile, o il suo calcolo è particolarmente costoso.

Si utilizza al posto di $\langle \cdot, \cdot \rangle$ il seguente prodotto scalare su $\mathbb{C}[z]$:

$$\langle \phi, \psi \rangle_{\star} := \frac{1}{2\pi i} \int_{\gamma} \phi(z) \psi(z) \frac{1}{f(z)} dz.$$

Alternativa *derivative-free*

Si può produrre una versione dell'algoritmo KSV che non richiede di valutare la derivata di f . Questo è utile nel caso in cui f' non è disponibile, o il suo calcolo è particolarmente costoso.

Si utilizza al posto di $\langle \cdot, \cdot \rangle$ il seguente prodotto scalare su $\mathbb{C}[z]$:

$$\langle \phi, \psi \rangle_{\star} := \frac{1}{2\pi i} \int_{\gamma} \phi(z) \psi(z) \frac{1}{f(z)} dz.$$

Anche per $\langle \cdot, \cdot \rangle_{\star}$ possiamo definire una successione di polinomi ortogonali formali $\{\phi_t^{\star}\}_{t=0}^{\infty}$.

Risulta che N è un indice regolare e che il polinomio $\phi_N^{\star}(z) := \prod_{k=1}^n (z - z_k)^{\nu_k}$

soddisfa le condizioni di ortogonalità

$$\langle z^p, \phi_N^{\star}(z) \rangle_{\star} = 0 \quad \forall p \in \mathbb{N},$$

- 1 Presentazione del problema
- 2 Risultati teorici
- 3 Algoritmo
- 4 Varianti
- 5 Interpolazione razionale**

Interpolazione razionale

L'algoritmo visto può essere interpretato in termini del seguente problema di interpolazione.

Interpolazione razionale

L'algoritmo visto può essere interpretato in termini del seguente problema di interpolazione.

Siano $t_\ell := \exp(2\ell\pi i/K)$ le radici K -esime dell'unità, con K intero positivo, e sia $g(z) := s_0 z^{K-1} + \dots + s_{K-1}$.

Supponiamo inoltre $g(t_\ell) \neq 0$ per ogni ℓ .

Interpolazione razionale

L'algoritmo visto può essere interpretato in termini del seguente problema di interpolazione.

Siano $t_\ell := \exp(2\ell\pi i/K)$ le radici K -esime dell'unità, con K intero positivo, e sia $g(z) := s_0 z^{K-1} + \dots + s_{K-1}$.

Supponiamo inoltre $g(t_\ell) \neq 0$ per ogni ℓ .

Cerchiamo due polinomi con $\deg p_\sigma \leq \sigma$ e $\deg q_\tau \leq \tau$, $\sigma + \tau = K - 1$, tali che

$$p_\sigma(t_\ell) - q_\tau(t_\ell)g(t_\ell) = 0, \quad \ell = 1, \dots, K. \quad (2)$$

La coppia (p_σ, q_τ) viene detta *multipoint Padé form* (MPF).

Interpolazione razionale

L'algoritmo visto può essere interpretato in termini del seguente problema di interpolazione.

Siano $t_\ell := \exp(2\ell\pi i/K)$ le radici K -esime dell'unità, con K intero positivo, e sia $g(z) := s_0 z^{K-1} + \dots + s_{K-1}$.

Supponiamo inoltre $g(t_\ell) \neq 0$ per ogni ℓ .

Cerchiamo due polinomi con $\deg p_\sigma \leq \sigma$ e $\deg q_\tau \leq \tau$, $\sigma + \tau = K - 1$, tali che

$$p_\sigma(t_\ell) - q_\tau(t_\ell)g(t_\ell) = 0, \quad \ell = 1, \dots, K. \quad (2)$$

La coppia (p_σ, q_τ) viene detta *multipoint Padé form* (MPF).

Osservazione

Se una MPF (p_σ, q_τ) è soluzione del problema (2), allora lo è anche la coppia $(\tilde{p}_\sigma, \tilde{q}_\tau)$ in cui sono stati semplificati i fattori comuni a p_σ e q_τ non corrispondenti alle t_ℓ .

Teorema

Sia (p_σ, q_τ) una MPF tale che $\text{g.c.d.}(p_\sigma, q_\tau)$ è un divisore di $z^K - 1$.
Se $K \geq 2n$, sono equivalenti:

- τ è un indice regolare per $\langle \cdot, \cdot \rangle$;
- $\deg q_\tau(z) = \tau$.

In questo caso vale anche $\phi_\tau \equiv q_\tau$ e l'indice regolare successivo a τ è $K - \deg p_\sigma(z)$.

Teorema

Sia (p_σ, q_τ) una MPF tale che $\text{g.c.d.}(p_\sigma, q_\tau)$ è un divisore di $z^K - 1$.
Se $K \geq 2n$, sono equivalenti:

- τ è un indice regolare per $\langle \cdot, \cdot \rangle$;
- $\deg q_\tau(z) = \tau$.

In questo caso vale anche $\phi_\tau \equiv q_\tau$ e l'indice regolare successivo a τ è $K - \deg p_\sigma(z)$.

Questo teorema ci permette di calcolare i FOP di $\langle \cdot, \cdot \rangle$ come denominatori di opportune MPF. I problemi di interpolazione possono essere risolti con un algoritmo di Bultheel e Van Barel [5].

Inoltre si possono calcolare le MPF corrispondenti a indici regolari in sequenza a partire da $\tau = 1$, controllando il grado di $p_\sigma(z)$ per determinare l'indice regolare successivo.

- Il controllo dei coefficienti di $p_\sigma(z)$ per determinarne il grado è analogo a quello della prima riga di un blocco di G_t : in effetti si può vedere che i coefficienti di $p_\sigma(z)$ sono i prodotti scalari di $q_\tau(z)$ con opportuni monomi z^P .

- Il controllo dei coefficienti di $p_\sigma(z)$ per determinarne il grado è analogo a quello della prima riga di un blocco di G_t : in effetti si può vedere che i coefficienti di $p_\sigma(z)$ sono i prodotti scalari di $q_\tau(z)$ con opportuni monomi z^P .
- Si verifica che l'indice n viene raggiunto al primo passo in cui avviene che $\deg p_\sigma(z) < \deg q_\tau(z)$. Questo ci fornisce un criterio d'arresto.

- Il controllo dei coefficienti di $p_\sigma(z)$ per determinarne il grado è analogo a quello della prima riga di un blocco di G_t : in effetti si può vedere che i coefficienti di $p_\sigma(z)$ sono i prodotti scalari di $q_\tau(z)$ con opportuni monomi z^P .
- Si verifica che l'indice n viene raggiunto al primo passo in cui avviene che $\deg p_\sigma(z) < \deg q_\tau(z)$. Questo ci fornisce un criterio d'arresto.
- Inoltre questo approccio basato sulle MPF ha le stesse buone proprietà dell'algoritmo originale per quanto riguarda i *clusters* di zeri.

P. Kravanja, T. Sakurai, and M. Van Barel, "On locating clusters of zeros of analytic functions," *BIT*, vol. 39, no. 4, pp. 646–682, 1999.

P. Kravanja and M. Van Barel, "A derivative-free algorithm for computing zeros of analytic functions," *Computing*, vol. 63, no. 1, pp. 69–91, 1999.

T. Sakurai, P. Kravanja, H. Sugiura, and M. Van Barel, "An error analysis of two related quadrature methods for computing zeros of analytic functions," in *Proceedings of the International Conference on Recent Advances in Computational Mathematics (ICRACM 2001) (Matsuyama)*, vol. 152, pp. 467–480, 2003.

I. Gohberg and I. Koltracht, "Mixed, componentwise, and structured condition numbers," *SIAM J. Matrix Anal. Appl.*, vol. 14, no. 3, pp. 688–704, 1993.

A. Bultheel and M. Van Barel, *Linear algebra, rational approximation and orthogonal polynomials*, vol. 6 of *Studies in Computational Mathematics*. North-Holland Publishing Co., Amsterdam, 1997.