



UNIVERSITÀ DEGLI STUDI DI CAGLIARI
FACOLTÀ DI SCIENZE
Corso di laurea Magistrale in Matematica

Problemi di deconvoluzione e deblurring:
analisi e determinazione della soluzione con
alcuni metodi di regolarizzazione

Relatore:

Prof. Giuseppe Rodriguez

Laureando:

Marco Boi

mat. 60/65/65118

Anno accademico 2023/2024

24 Settembre 2024

Ai miei amati genitori

Indice

Introduzione	5
1 Nozioni preliminari	7
1.1 La trasformata di Fourier discreta e la FFT	7
1.1.1 Definizione e proprietà	7
1.1.2 Forma matriciale della DFT e della IDFT	8
1.1.3 Trasformata discreta dei coseni (DCT) e trasformata discreta dei seni (DST)	9
1.1.4 DFT della convoluzione discreta	10
1.1.5 La Fast Fourier Transform	11
1.1.6 La DFT bidimensionale	12
1.2 La decomposizione a valori singolari	12
1.2.1 Valori singolari e vettori singolari	12
1.2.2 Approssimazioni di matrici e definizione di δ -rango di una matrice	13
2 Problemi di deconvoluzione	15
2.1 Equazioni integrali di Fredholm del primo tipo: definizione ed esempi	15
2.2 Discretizzazione dell'equazione integrale di Fredholm di prima specie	17
2.3 Deconvoluzione in una dimensione	19
2.3.1 Matrici di Toeplitz e loro SVD	20
2.3.2 Matrici circolanti e convoluzione	23
2.3.3 Moltiplicazione matrice-vettore attraverso la FFT	25
2.4 Deconvoluzione in due dimensioni	26
2.4.1 Il prodotto di Kronecker e la "vec notation"	26
2.4.2 Discretizzazione per kernel con variabili separabili	28
3 Image restoration	31
3.1 Deduzione dell'equazione integrale di Fredholm associata al deblurring	31
3.2 Il problema dell'image deconvolution	33

3.2.1	Formulazione del problema nel continuo	34
3.2.2	Formulazione del problema nel discreto	36
3.2.3	Esempi di blurring	36
	Linear motion blur	36
	Sfocatura da non messa a fuoco	37
	Sfocatura originata da turbolenza atmosferica	37
3.3	Boundary conditions	38
3.3.1	Boundary conditions per immagini in una dimensione	38
	Zero boundary condition	40
	Periodic boundary condition	40
	Reflexive boundary condition	41
3.3.2	Boundary conditions per immagini bidimensionali	42
3.4	Proprietà numeriche del problema dell'immagine deconvolution	44
3.4.1	Mal posizione del problema nel continuo	45
3.4.2	Discretizzazione e mal condizionamento	47
4	Metodi di regolarizzazione	49
4.1	I metodi di filtraggio spettrale	49
4.1.1	La SVD troncata (TSVD)	51
4.1.2	Regolarizzazione di Tikhonov	52
4.1.3	TSVD e regolarizzazione di Tikhonov per problemi bidimensionali	54
4.2	Criteri per la scelta del parametro di regolarizzazione	55
	Il principio di discrepanza	55
	Generalized cross-validation (GCV)	55
	Metodo della L-curva	56
4.3	Metodi iterativi di regolarizzazione	57
4.3.1	Metodo del gradiente	57
4.3.2	Metodo del gradiente coniugato	59
4.3.3	Metodo del gradiente coniugato per problemi ai minimi quadrati (CGLS)	60
4.3.4	CGLS per problemi in dimensione 2	61
5	Test numerici e sperimentazioni	63
5.1	I dati del problema: costruzione dell'immagine sfocata	64
5.2	Determinazione della soluzione numerica mediante metodi di regolarizzazione diretti	66
5.2.1	Decomposizione di Kronecker	67
5.2.2	Regolarizzazione tramite la TSVD	69

<i>INDICE</i>	4
Scelta del parametro usando la GCV	70
Scelta del parametro usando il metodo della L-curva	71
Scelta del parametro usando il principio di discrepanza	73
Confronto tra le soluzioni	75
5.2.3 Regularizzazione tramite filtraggio di Tikhonov	76
5.3 Determinazione della soluzione numerica mediante metodi di regola- rizzazione iterativi: soluzione per CGLS	77
Conclusioni	80
Bibliografia	81

Introduzione

Quando viene scattata una fotografia vorremmo fare in modo che essa sia una rappresentazione più fedele possibile di ciò che osserviamo. Tuttavia, per quanto ci si impegni nel cercare di essere precisi, ogni immagine digitalizzata è affetta da errore, ed è sempre più o meno *sfocata*. Le ragioni per cui si origina la sfocatura possono essere legate, ad esempio, a una non corretta messa a fuoco della lente, alla presenza di turbolenza nell'atmosfera, alla presenza di oggetti in movimento sullo sfondo o a qualsiasi altro dettaglio, più o meno evidente, che ostacoli il percorso dei raggi di luce diretti allo strumento ottico scelto per la ripresa fotografica.

In analisi dei segnali, con l'espressione di *image deblurring* si intende l'insieme di tutti i tentativi atti al ripristino dell'immagine nitida attraverso l'utilizzo di modelli matematici idonei: in analisi numerica il problema del deblurring si traduce, dunque, nella ricerca di algoritmi efficienti e stabili che consentano di recuperare quanta più informazione possibile dai dati a disposizione.

Da un punto di vista più generale, il deblurring appartiene alla classe dei cosiddetti problemi di deconvoluzione, i quali godono di numerose applicazioni non soltanto nell'ambito dell'elaborazione digitale delle immagini e nell'analisi dei segnali, ma anche in altre scienze applicate.

Lo scopo di questo elaborato è quello di approfondire gli aspetti matematici e numerici dell'immagine deblurring per immagini in bianco e nero, giustificando, inizialmente, l'appartenenza del fenomeno alla classe dei problemi di deconvoluzione, per poi analizzare la particolare struttura matriciale che il problema può assumere una volta discretizzato in funzione della particolare condizione imposta sui contorni dell'immagine (*boundary condition*).

La proprietà numerica del mal condizionamento, che vedremo essere caratteristica per i problemi discreti derivanti da problemi di deconvoluzione e in particolare per il deblurring, apre la questione legata alla *regolarizzazione* della soluzione al problema. In questo lavoro verranno analizzati e sottoposti a test numerici alcuni metodi di regolarizzazione: il *metodo della SVD troncata (TSVD)*, il *metodo di regolarizzazione di Tikhonov* e il metodo iterativo del *gradiente coniugato per problemi ai minimi quadrati*.

Nell'ambito dei metodi di regolarizzazione diretti (TSVD e metodo di Tikhonov) verrà inoltre approfondito il problema legato alla scelta del parametro di regolarizzazione, per il quale verranno presentati e testati tre differenti criteri di scelta.

La possibilità di adottare diversi metodi numerici per la determinazione della soluzione al problema consentirà di poter confrontare tra loro diversi risultati, nonché di avere una più accurata conoscenza del problema stesso, soprattutto secondo il punto di vista dell'Analisi Numerica.

Capitolo 1

Nozioni preliminari

1.1 La trasformata di Fourier discreta e la FFT

1.1.1 Definizione e proprietà

La *trasformata di Fourier discreta* (DFT) può essere definita come l'operatore

$$DFT : \mathbb{C}^N \longrightarrow \mathbb{C}^N$$

tale che

$$\left\{ f_n \right\}_{n=0}^{N-1} \mapsto \left\{ \hat{f}_n \right\}_{n=0}^{N-1},$$

dove

$$\hat{f}_k = \frac{1}{N} \sum_{n=0}^{N-1} f_n \omega_N^{-nk}, \quad k = 0, \dots, N-1, \quad (1.1)$$

con $\omega_N = e^{\frac{2\pi i}{N}}$ radice N -esima dell'unità.

Proprietà. (ortogonalità dell'esponenziale complesso)

$$\sum_{k=0}^{N-1} \omega_N^{rk} \omega_N^{-sk} = \begin{cases} N & \text{se } r = s \\ 0 & \text{se } r \neq s \end{cases}$$

Dimostrazione.

$$\sum_{k=0}^{N-1} \omega_N^{rk} \omega_N^{-sk} = \sum_{k=0}^{N-1} (\omega_N^{r-s})^k = \begin{cases} \frac{1 - (\omega_N^{r-s})^N}{1 - \omega_N^{r-s}} = 0, & r \neq s \\ N, & r = s \end{cases}$$

□

Utilizzando la proprietà di ortogonalità dell'esponenziale complesso è possibile definire la trasformata di Fourier inversa $DFT^{-1} = IDFT$ come

$$f_n = \sum_{k=0}^{N-1} \hat{f}_k \omega_N^{nk}, \quad n = 0, 1, \dots, N-1. \quad (1.2)$$

La IDFT così definita rappresenta proprio l'inversa della DFT, infatti:

$$\sum_{k=0}^{N-1} \hat{f}_k \omega_N^{nk} = \sum_{k=0}^{N-1} \left[\frac{1}{N} \sum_{r=0}^{N-1} f_r \omega_N^{-rk} \right] \omega_N^{nk} = \sum_{r=0}^{N-1} \left[\frac{1}{N} \sum_{k=0}^{N-1} \omega_N^{nk} \omega_N^{-rk} \right] f_r$$

Per la proprietà dell'esponenziale complesso si ha allora che:

$$\sum_{k=0}^{N-1} \hat{f}_k \omega_N^{nk} = \sum_{r=0}^{N-1} \left[\frac{1}{N} N \delta_{nr} \right] f_r = f_n.$$

1.1.2 Forma matriciale della DFT e della IDFT

Siano $f = \begin{bmatrix} f_0 \\ f_1 \\ \vdots \\ f_{N-1} \end{bmatrix}$ e $\hat{f} = \begin{bmatrix} \hat{f}_0 \\ \hat{f}_1 \\ \vdots \\ \hat{f}_{N-1} \end{bmatrix}$ i vettori di dimensione N rappresentanti una sequenza di N complessi e la loro *DFT*.

Indicando con

$$\mathcal{F} = \begin{bmatrix} \omega_N^0 & \omega_N^0 & \omega_N^0 & \dots & \omega_N^0 \\ \omega_N^0 & \omega_N^{-1} & \omega_N^{-2} & \dots & \omega_N^{-(N-1)} \\ \omega_N^0 & \omega_N^{-2} & \omega_N^{-4} & \dots & \omega_N^{-2(N-1)} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \omega_N^0 & \omega_N^{-(N-1)} & \omega_N^{-2(N-1)} & \dots & \omega_N^{-(N-1)(N-1)} \end{bmatrix} \quad (1.3)$$

la matrice di Fourier e con

$$\mathcal{F}^* = \begin{bmatrix} \omega_N^0 & \omega_N^0 & \omega_N^0 & \dots & \omega_N^0 \\ \omega_N^0 & \omega_N^1 & \omega_N^2 & \dots & \omega_N^{(N-1)} \\ \omega_N^0 & \omega_N^2 & \omega_N^4 & \dots & \omega_N^{2(N-1)} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \omega_N^0 & \omega_N^{(N-1)} & \omega_N^{2(N-1)} & \dots & \omega_N^{(N-1)(N-1)} \end{bmatrix} \quad (1.4)$$

possiamo esprimere la *DFT* e la *IDFT* utilizzando la notazione matriciale:

$$\hat{f} = \frac{1}{N} \mathcal{F} f, \quad (1.5)$$

$$f = \mathcal{F}^* \hat{f}. \quad (1.6)$$

Osservazione. Utilizzando la proprietà di ortogonalità dell'esponenziale complesso si ottiene che

$$\mathcal{F} \mathcal{F}^* = N I.$$

Infatti:

$$[\mathcal{F} \mathcal{F}^*]_{ij} = \sum_{k=0}^{N-1} \omega_N^{-ik} \omega_N^{kj} = N \delta_{ij}$$

1.1.3 Trasformata discreta dei coseni (DCT) e trasformata discreta dei seni (DST)

Se il segnale in input gode di proprietà di simmetria allora il calcolo della DFT può essere semplificato e conduce alla definizione di altre trasformate.

1. $f_n = f_{-n}$ (*simmetria pari*).

$$\begin{aligned}
 \hat{f}_k &= \frac{1}{N} \sum_{n=-\frac{N}{2}}^{\frac{N}{2}-1} f_n \omega_N^{-nk} = \frac{1}{N} \left[f_0 + \sum_{n=1}^{\frac{N}{2}-1} \left(f_n \omega_N^{-nk} + f_{-n} \omega_N^{nk} \right) + f_{-\frac{N}{2}} \omega_N^{k\frac{N}{2}} \right] = \\
 &= \frac{1}{N} \left[f_0 + \sum_{n=1}^{\frac{N}{2}-1} f_n \left(e^{-\frac{2\pi ink}{N}} + e^{\frac{2\pi ink}{N}} \right) + f_{\frac{N}{2}} (-1)^k \right] = \\
 &= \frac{1}{N} \left[f_0 + 2 \sum_{n=1}^{\frac{N}{2}-1} f_n \cos \left(\frac{2\pi nk}{N} \right) + f_{\frac{N}{2}} \cos(\pi k) \right] = \\
 &= \frac{1}{N/2} \left[\frac{1}{2} f_0 + \sum_{n=1}^{\frac{N}{2}-1} f_n \cos \left(\frac{\pi nk}{N/2} \right) + \frac{1}{2} f_{\frac{N}{2}} \cos(\pi k) \right]
 \end{aligned}$$

Si definisce, dunque, la *trasformata discreta dei coseni (DCT)*:

$$F_k = \frac{1}{N} \sum_{n=0}^N {}^{\prime\prime} f_n \cos \left(\frac{\pi nk}{N} \right), \quad (1.7)$$

dove ${}^{\prime\prime}$ indica che il primo e l'ultimo termine hanno peso $\frac{1}{2}$.

2. $f_n = -f_{-n}$ (*simmetria dispari*).

$$\begin{aligned}
 \hat{f}_k &= \frac{1}{N} \sum_{n=-\frac{N}{2}}^{\frac{N}{2}-1} f_n \omega_N^{-nk} = \frac{1}{N} \left[f_0 + \sum_{n=1}^{\frac{N}{2}-1} \left(f_n \omega_N^{-nk} + f_{-n} \omega_N^{nk} \right) + f_{-\frac{N}{2}} \omega_N^{k\frac{N}{2}} \right] = \\
 &= \frac{1}{N} \left[\sum_{n=1}^{\frac{N}{2}-1} f_n \left(e^{-\frac{2\pi ink}{N}} - e^{\frac{2\pi ink}{N}} \right) \right] = \frac{1}{N} \left[-2i \sum_{n=1}^{\frac{N}{2}-1} f_n \sin \left(\frac{2\pi nk}{N} \right) \right] = \\
 &= -\frac{i}{N/2} \left[\sum_{n=1}^{\frac{N}{2}-1} f_n \sin \left(\frac{\pi nk}{N/2} \right) \right]
 \end{aligned}$$

Si definisce, dunque, la *trasformata discreta dei seni (DST)*:

$$F_k = \frac{1}{N} \sum_{n=1}^{N-1} f_n \sin \left(\frac{\pi nk}{N} \right). \quad (1.8)$$

3. $f_n \in \mathbb{R}$ (simmetria reale).

In questo caso si ottiene la real DFT (RDFT), che è equivalente a una coppia DCT/DST:

$$\hat{f}_k = \frac{1}{N} \sum_{n=-\frac{N}{2}}^{\frac{N}{2}-1} f_n \left[\cos\left(\frac{2\pi nk}{N}\right) - i \sin\left(\frac{2\pi nk}{N}\right) \right] = \Re(\hat{f}_k) + i\Im(\hat{f}_k),$$

$$\text{avendo posto } \Re(\hat{f}_k) = \frac{1}{N} \sum_{n=-\frac{N}{2}}^{\frac{N}{2}-1} f_n \cos\left(\frac{2\pi nk}{N}\right) \text{ e } \Im(\hat{f}_k) = -\frac{1}{N} \sum_{n=-\frac{N}{2}}^{\frac{N}{2}-1} f_n \sin\left(\frac{2\pi nk}{N}\right)$$

1.1.4 DFT della convoluzione discreta

Teorema. Siano f e g due sequenze N -periodiche (cioè tali per cui $f_{j\pm N} = f_j$, $g_{j\pm N} = g_j$) e siano F_k e G_k la k -esima componente delle rispettive DFT. Allora la DFT della convoluzione discreta $h = f * g$ ha componenti

$$H_k = DFT\{f * g\}_k = NF_k G_k, \quad k = 1, \dots, n. \quad (1.9)$$

Dimostrazione.

$$h_n = f_n * g_n = \sum_{j=0}^{N-1} f_j g_{(n-j) \bmod n}.$$

Usando l'inversa della trasformata di Fourier discreta (1.2) si deduce che

$$\begin{aligned} h_n &= \sum_{j=0}^{N-1} \left(\sum_{p=0}^{N-1} F_p \omega_N^{jp} \left(\sum_{k=0}^{N-1} G_k \omega_N^{(n-j)k} \right) \right) = \sum_{k=0}^{N-1} G_k \omega_N^{nk} \sum_{p=0}^{N-1} F_p \sum_{j=0}^{N-1} \omega_N^{j(p-k)} = \\ &= \sum_{k=0}^{N-1} G_k \omega_N^{nk} \sum_{p=0}^{N-1} F_p N \hat{\delta}_N(p-k), \end{aligned}$$

$$\text{dove } \hat{\delta}_N(x) = \begin{cases} 1 & \text{se } x \equiv_N 0 \\ 0 & \text{altrimenti} \end{cases}.$$

Pertanto

$$h_n = \sum_{k=0}^{N-1} G_k F_k N \omega_N^{nk} = N \sum_{k=0}^{N-1} F_k G_k \omega_N^{nk} = N \cdot IDFT\{F_k G_k\}_n$$

da cui segue che

$$DFT\{h_n\}_k = DFT\{f_n * g_n\}_k = NF_k G_k.$$

□

1.1.5 La Fast Fourier Transform

Il calcolo della DFT così come è stata definita nella (1.1) ha una complessità computazionale di $\mathcal{O}(n^2)$. Tuttavia, nelle applicazioni numeriche si privilegiano strategie di calcolo alternative per abbattarne il costo computazionale. Per questo motivo ha senso parlare di algoritmi ottimizzati che implementano la cosiddetta *Fast Fourier Transform (FFT)*.

Il più diffuso tra questi è l'*algoritmo di Cooley-Tukey* (1965) che utilizza il cosiddetto “metodo di splitting”, di seguito descritto.

Sia $N = 2^M$ e sia

$$X_k = \sum_{n=0}^{N-1} x_n \omega_N^{-nk} = \sum_{n=0}^{\frac{N}{2}-1} (x_{2n} \omega_N^{-nk} + x_{2n+1} \omega_N^{-(2n+1)k}), \quad k = 0, 1, \dots, N-1.$$

Ponendo $y_n = x_{2n}$ e $z_n = x_{2n+1}$ si ottiene

$$X_k = \sum_{n=0}^{\frac{N}{2}-1} y_n \omega_N^{-2nk} + \omega_N^{-k} \sum_{n=0}^{\frac{N}{2}-1} z_n \omega_N^{-2nk}. \quad (1.10)$$

Osservando che

$$\omega_N^{-2nk} = e^{-\frac{4\pi ink}{N}} = e^{-\frac{2\pi ink}{(N/2)}} = \omega_{\frac{N}{2}}^{-nk},$$

la (1.10) può essere riscritta come

$$X_k = \sum_{n=0}^{\frac{N}{2}-1} y_n \omega_{\frac{N}{2}}^{-nk} + \omega_N^{-k} \sum_{n=0}^{\frac{N}{2}-1} z_n \omega_{\frac{N}{2}}^{-nk} = Y_k + \omega_N^{-k} Z_k.$$

In questo modo, la DFT originale, avente lunghezza N è stata espressa attraverso una semplice combinazione di due DFT aventi entrambe lunghezza $\frac{N}{2}$.

Per $k = 0, \dots, \frac{N}{2} - 1$ valgono dunque le seguenti relazioni:

$$\begin{cases} X_k = Y_k + \omega_N^{-k} Z_k \\ X_{k+\frac{N}{2}} = Y_{k+\frac{N}{2}} + \omega_N^{-(k+\frac{N}{2})} Z_{k+\frac{N}{2}} \end{cases}.$$

Infine, sfruttando il fatto che Y_k e Z_k sono $\frac{N}{2}$ -periodiche, le precedenti risultano equivalenti alle

$$\begin{cases} X_k = Y_k + \omega_N^{-k} Z_k \\ X_{k+\frac{N}{2}} = Y_k - \omega_N^{-k} Z_k \end{cases}, \quad k = 0, \dots, \frac{N}{2} - 1, \quad (1.11)$$

che vengono anche dette *formule di combinazione* o *relazioni “butterfly”*.

Ripetendo il procedimento, le due DFT di lunghezza $\frac{N}{2}$ vengono decomposte nella somma di due *DFT* di lunghezza $\frac{N}{4} = \frac{N}{2^2}$; ciascuna di queste, a sua volta, verrà

espressa come combinazione di due DFT di lunghezza $\frac{N}{2^3}$ e così via. Alla M -esima iterazione, essendo $M = \log_2 N$, la DFT di partenza viene dunque espressa come somma di N DFT di lunghezza 1.

Osserviamo, inoltre, che a ogni passo le (1.11) vengono applicate $\frac{N}{2}$ volte e richiedono una moltiplicazione e due addizioni complesse. La complessità computazionale totale risulta pertanto essere $N \log_2 N$.

1.1.6 La DFT bidimensionale

Sia f la funzione che descrive un segnale bidimensionale e definita nel dominio

$$\left\{ (x, y) \text{ tali che } -\frac{A}{2} \leq x \leq \frac{A}{2}, \quad -\frac{B}{2} \leq y \leq \frac{B}{2} \right\}$$

discretizzato lungo le direzioni x e y attraverso le suddivisioni aventi i seguenti passi:

$$\Delta x = \frac{A}{M}, \quad \Delta y = \frac{B}{N}.$$

Il segnale f può dunque essere calcolato nei punti delle griglia $(x_m, y_n) = (m\Delta x, n\Delta y)$, per $n = 0, \dots, N-1$, $m = 0, \dots, M-1$.

In analogia con la (1.1) definiamo la DFT *bidimensionale* di $f_{mn} = f(x_m, y_n)$ come

$$DFT2\{f_{mn}\}_{jk} = \frac{1}{MN} \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} f_{mn} \omega_M^{-mj} \omega_N^{-nk}, \quad j = 0, \dots, M-1, \quad k = 0, \dots, N-1. \quad (1.12)$$

Si dimostra che l'inversa della (1.12) è espressa dalla seguente espressione:

$$IDFT2\{F_{jk}\}_{mn} = \sum_{j=0}^{M-1} \sum_{k=0}^{N-1} F_{jk} \omega_M^{mj} \omega_N^{nk}, \quad m = 0, \dots, M-1, \quad n = 0, \dots, N-1. \quad (1.13)$$

In maniera analoga a quanto fatto nel caso unidimensionale è possibile, inoltre, definire le trasformate DCT2 e DST2.

1.2 La decomposizione a valori singolari

1.2.1 Valori singolari e vettori singolari

Teorema. Sia $A \in \mathbb{C}^{m \times n}$ una matrice di rango r . Allora esistono due matrici unitarie $U \in \mathbb{C}^{m \times m}$ e $V \in \mathbb{C}^{n \times n}$ tali che

$$A = U \Sigma V^H, \quad \Sigma = \begin{bmatrix} \Sigma_1 & 0 \\ 0 & 0 \end{bmatrix}$$

dove V^H indica la matrice trasposta coniugata della matrice V , $\Sigma \in \mathbb{R}^{m \times n}$, $\Sigma_1 = \text{diag}(\sigma_1, \sigma_2, \dots, \sigma_r)$ e

$$\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_r > 0.$$

I valori σ_i sono detti valori singolari di A e i vettori colonna delle matrici

$$U = \begin{bmatrix} \mathbf{u}_1 & \dots & \mathbf{u}_m \end{bmatrix}, \quad V = \begin{bmatrix} \mathbf{v}_1 & \dots & \mathbf{v}_n \end{bmatrix}$$

sono detti rispettivamente vettori singolari destri e sinistri associati a σ_i , per $i = 1, \dots, r$.

Osservazione. La SVD di una matrice A di rango r può essere scritta come la somma di r matrici di rango 1:

$$A = U_1 \Sigma_1 V_1^H = \sum_{i=1}^r \sigma_i u_i v_i^H, \quad (1.14)$$

dove

$$U_1 = (u_1, \dots, u_r) \quad V_1 = (v_1, \dots, v_r).$$

1.2.2 Approssimazioni di matrici e definizione di δ -rango di una matrice

Teorema. Sia $A \in \mathbb{C}^{m \times n}$ matrice di rango r e SVD

$$A = U \Sigma V^H = \sum_{i=1}^r \sigma_i u_i v_i^H.$$

Sia $\mathcal{M}_k^{m \times n}(\mathbb{C})$ l'insieme delle matrici di rango $k < r$. Allora

$$\min_{X \in \mathcal{M}_k^{m \times n}(\mathbb{C})} \|A - X\|_2 \quad (1.15)$$

è raggiunto per $X = B$, dove

$$B = \sum_{i=1}^k \sigma_i u_i v_i^H, \quad \|A - B\|_2 = \sigma_{k+1}.$$

Osservazione. Come caso particolare del teorema, osserviamo che se $\text{rank}(A) = n$ allora il valore singolare σ_n realizza la minima distanza, rispetto alla norma-2, tra A e l'insieme di tutte le matrici singolari.

Osservazione. Il teorema può essere generalizzato anche al caso in cui la norma in considerazione sia la norma di Frobenius. In questo caso la minima distanza è data da

$$\|A - B\|_F = \sqrt{\sigma_{k+1}^2 + \dots + \sigma_r^2}$$

Definizione. Diciamo che una matrice $A \in \mathbb{C}^{m \times n}$ ha δ -rango pari a k se

$$k = \min\{\text{rank}(B) \mid \|A - B\|_2 \leq \delta\} \quad (1.16)$$

Da questa definizione e dal teorema precedente, segue che, se $k < n$ allora

$$\inf_{\text{rank}(B) \leq k} \|A - B\|_2 = \sigma_{k+1},$$

dove σ_i , $i = 1, 2, \dots, \min(m, n)$, sono i valori singolari di A .

Tale inf è raggiunto dalla matrice

$$B = \sum_{i=1}^k \sigma_i u_i v_i^t.$$

Pertanto, si può affermare, equivalentemente, che la matrice A ha δ -rango pari a k se e solo se

$$\sigma_1 \geq \dots \geq \sigma_k > \delta \geq \sigma_{k+1} \geq \dots \quad (1.17)$$

Capitolo 2

Problemi di deconvoluzione

2.1 Equazioni integrali di Fredholm del primo tipo: definizione ed esempi

Il problema del deblurring rientra nella categoria dei problemi inversi, cioè quei problemi in cui partendo da alcune misurazioni di dati si intende effettuare delle inferenze su delle grandezze o quantità incognite che non sono direttamente accessibili attraverso le misurazioni.

Oltre all'immagine deblurring (termine dall'inglese *"to blur"*, cioè appannarsi, offuscarsi), vi sono altri numerosi esempi di problemi inversi che si presentano nei diversi ambiti delle scienze applicate:

- in *sismologia*, il cui obiettivo è quello di rappresentare i vari strati della terra partendo dalla ricezione, sulla superficie terrestre, del segnale prodotto dalle onde sismiche;
- nell'ambito della radiodiagnostica, per esempio, la *tomografia computerizzata (CT)* riproduce le immagini in sezione e tridimensionali dell'anatomia di un organo del corpo umano note le misure dell'attenuazione di un fascio di raggi-X che attraversa l'organo di interesse.

Gli esempi citati, l'immagine deblurring, e tanti altri problemi inversi, possono essere formulati utilizzando un'opportuna equazione integrale, detta *equazione integrale di Fredholm del primo tipo*.

Definizione. Si dice *equazione integrale lineare di Fredholm del primo tipo* un'equazione della forma

$$g(x) = \int_a^b K(x, y)f(y)dy , \quad (2.1)$$

in cui

1. la funzione $K(x, y)$, detta *kernel*, è nota in $\{(x, y) : a \leq x, y \leq b\}$
2. la funzione $g(x)$ è nota per $a \leq x \leq b$.

Esempio 1 (la trasformata di Laplace). La formula di inversione della trasformata di Laplace è equivalente alla soluzione di un'equazione integrale lineare di Fredholm del primo tipo.

Infatti, sia

$$g(x) = \int_0^{\infty} e^{-xy} f(y) dy \quad (2.2)$$

l'integrale di Fredholm di prima specie che esprime la trasformata di Laplace di una certa funzione f , in cui g è una data funzione analitica nel semipiano complesso $\sigma = \{x \in \mathbb{C} : \operatorname{Re}\{x\} > \sigma[f]\}$, intendendo con $\sigma[f] := \inf\{\operatorname{Re}(x) : e^{-xy} f(y) \in L^1(\mathbb{R}^+)\}$, e f è la funzione incognita da determinare.

Allora, per la formula di inversione della trasformata di Laplace, sappiamo che la funzione

$$f(y) = \frac{1}{2\pi i} \int_{c-i\infty}^{c+i\infty} e^{yx} g(x) dx, \quad (2.3)$$

costituisce la soluzione dell'equazione integrale (2.2).

Osserviamo, tuttavia, che tale soluzione non è del tutto soddisfacente, in quanto nella (2.2) $g(x)$ è nota in \mathbb{R} , mentre nella (2.3) è nota in un semipiano del piano complesso \mathbb{C} . Potrebbero, allora, essere usate altre formule di inversione numeriche e reali, come ad esempio la seguente:

$$f(y) = \lim_{k \rightarrow \infty} \left[\frac{(-1)^k}{k!} g^{(k)}\left(\frac{k}{y}\right) \right] \left(\frac{k}{y}\right)^{k+1},$$

la quale mostra che l'inversa della trasformata dipende dal comportamento della k -esima derivata di g , al tendere di k all'infinito (cfr. [7], pag. 8).

Esempio 2 (Deconvoluzione). Siano f e h due funzioni a valori reali e sia

$$g(s) := (f * h)(s) = \int_0^1 h(s-t) f(t) dt, \quad 0 \leq s \leq 1, \quad (2.4)$$

la loro *convoluzione* rispetto all'intervallo $[0, 1]$.

Ammettendo che h e g siano note, il problema generale di *deconvoluzione* consiste nella determinazione della funzione incognita f . Con queste ipotesi, l'equazione integrale (2.4) è una particolare equazione lineare di Fredholm del primo tipo, il cui kernel $K(s, t) = h(s-t)$ dipende solamente dalla differenza $s-t$ tra le due variabili.

Nella pratica, la funzione $g(s)$ è nota, attraverso delle misurazioni, in un insieme discreto di punti $\{s_1, \dots, s_n\}$. In generale, si può quindi affermare che le misure effettuate possono essere espresse nella forma

$$\tilde{g}(s_i) = g(s_i) + e_i, \quad i = 1, \dots, n,$$

dove le componenti del vettore $(e_1, \dots, e_n)^t$ indicano il rumore associato a ciascuna delle misurazioni.

2.2 Discretizzazione dell'equazione integrale di Fredholm di prima specie

Le regole di quadratura vengono utilizzate per approssimare un generico integrale definito:

$$\int_a^b \phi(t) dt \approx \sum_{j=1}^n w_j \phi(t_j) ,$$

dove t_1, \dots, t_n sono i nodi per la particolare regola di quadratura e w_1, \dots, w_n sono i pesi corrispondenti al metodo di quadratura adottato.

Pertanto, una regola di quadratura consente di approssimare l'integrale che compare nell'equazione integrale di Fredholm (2.1):

$$\int_0^1 K(s, t) f(t) dt \approx \sum_{j=1}^n w_j K(s, t_j) \tilde{f}(t_j) = \psi(s) ,$$

dove \tilde{f} indica la perturbazione di f . Richiedendo che

$$\psi(s_i) = g(s_i), \quad i = 1, \dots, n$$

l'equazione integrale di Fredholm può essere riscritta in forma matriciale

$$A\mathbf{x} = \mathbf{b}, \tag{2.5}$$

dove

$$A = \begin{bmatrix} w_1 K(s_1, t_1) & w_2 K(s_1, t_2) & \dots & w_n K(s_1, t_n) \\ w_1 K(s_2, t_1) & w_2 K(s_2, t_2) & \dots & w_n K(s_2, t_n) \\ \vdots & \vdots & & \vdots \\ w_1 K(s_n, t_1) & w_2 K(s_n, t_2) & \dots & w_n K(s_n, t_n) \end{bmatrix}, \quad x = \begin{bmatrix} \tilde{f}(t_1) \\ \tilde{f}(t_2) \\ \vdots \\ \tilde{f}(t_n) \end{bmatrix}, \quad b = \begin{bmatrix} g(s_1) \\ g(s_2) \\ \vdots \\ g(s_n) \end{bmatrix} .$$

Nella pratica si verifica spesso che la matrice A in (2.5) è mal condizionata. Pertanto, la soluzione $\mathbf{x}_{\text{naive}} = A^{-1}\mathbf{b}$ è in realtà molto distante dalla soluzione esatta.

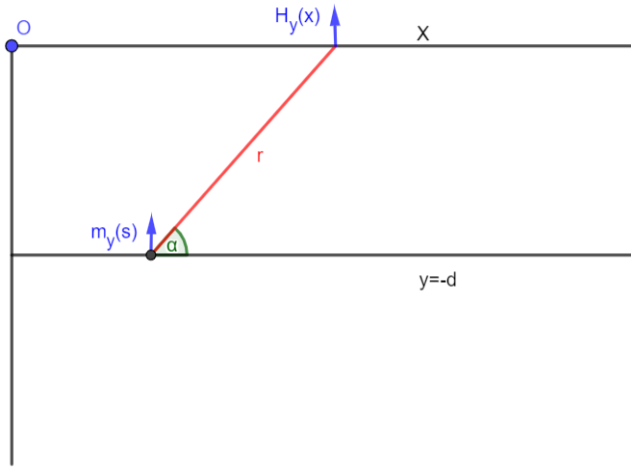
Esempio 3. Supponiamo che sia stata determinata, attraverso degli scavi, la presenza di un deposito minerale in uno strato della terra che dista dalla superficie terrestre di una nota distanza d . Il deposito produce un campo magnetico, misurabile sulla superficie attraverso un magnetometro, e la cui intensità alla superficie è un indicatore della quantità di minerale presente. Si vogliono determinare le zone in cui è presente la maggior quantità di minerale, cioè quelle in cui è più conveniente

effettuare lo scavo.

Il problema può essere formulato adottando un modello semplificato in una dimensione.

Consideriamo un riferimento cartesiano $R(O, xy)$ rispetto al quale:

1. la retta $y = -d$ determina la posizione dello strato;
2. nel generico punto $(s, -d)$, dove $0 \leq s \leq 1$, la componente verticale del campo magnetico ha intensità incognita $m_y(s)$;
3. in un generico punto $P(x, 0)$ sull'asse x è misurata la componente verticale dell'intensità del campo magnetico $H_y(x)$.



In $P(x, 0)$ il contributo a $H_y(x)$ dell'intensità $m_y(s)$ del campo magnetico prodotto dallo strato infinitesimo ds è noto essere inversamente proporzionale al quadrato della distanza r tra i punti $(s, -d)$ e P , cioè, con riferimento alla figura:

$$dH_y(x) = \frac{\sin(\alpha)m_y(s)ds}{r^2} = \frac{\sin(\alpha)m_y(s)ds}{(x-s)^2 + d^2} = \frac{d \cdot m_y(s)ds}{[(x-s)^2 + d^2]^{3/2}}$$

da cui si giunge all'*equazione integrale lineare di Fredholm del primo tipo*

$$H_y(x) = \int_0^1 \frac{d \cdot m_y(s)ds}{[(x-s)^2 + d^2]^{3/2}}, \quad (2.6)$$

il cui kernel $K(x, s) = h(x-s) = \frac{d}{[(x-s)^2 + d^2]^{3/2}}$.

Approssimiamo l'integrale in (2.6) utilizzando, per semplicità, la *formula di quadratura del punto medio* e partizioniamo l'intervallo $[0, 1]$ in n sottointervalli di uguale ampiezza. Effettuando questa scelta, risulta che:

$$\int_0^1 \frac{d \cdot m_y(s)ds}{[(x-s)^2 + d^2]^{3/2}} = \int_0^1 h(x-s)m_y(s)ds \approx \sum_{j=1}^n \frac{1}{n} h\left(x - \frac{j-0.5}{n}\right) m_y\left(\frac{j-0.5}{n}\right) = \psi(x).$$

Imponendo le condizioni di interpolazione

$$\psi(x_i) = H_y(x_i), \quad i = 1, \dots, n,$$

si deduce che la matrice del sistema è la seguente:

$$A = \begin{bmatrix} \frac{d/n}{[(x_1-s_1)^2+d^2]^{3/2}} & \frac{d/n}{[(x_1-s_2)^2+d^2]^{3/2}} & \cdots & \frac{d/n}{[(x_1-s_n)^2+d^2]^{3/2}} \\ \frac{d/n}{[(x_2-s_1)^2+d^2]^{3/2}} & \frac{d/n}{[(x_2-s_2)^2+d^2]^{3/2}} & \cdots & \frac{d/n}{[(x_2-s_n)^2+d^2]^{3/2}} \\ \vdots & \vdots & & \vdots \\ \frac{d/n}{[(x_n-s_1)^2+d^2]^{3/2}} & \frac{d/n}{[(x_n-s_2)^2+d^2]^{3/2}} & \cdots & \frac{d/n}{[(x_n-s_n)^2+d^2]^{3/2}} \end{bmatrix}.$$

Considerando poi che $x_i = s_i = \frac{i-0.5}{n}$ otteniamo che

$$[A]_{ij} = [A]_{ji} = \frac{dn^2}{[(i-j)^2 + n^2d^2]^{3/2}}.$$

A questo punto, attribuendo al dato relativo alla profondità dei valori arbitrari è possibile constatare sperimentalmente che, indipendentemente dal valore di n , la matrice del sistema è fortemente *mal condizionata*.

Per $d = 0.25$, per esempio, si ottengono i seguenti risultati:

n	$\text{cond}(A)$
20	$3.7714 \cdot 10^5$
50	$3.3774 \cdot 10^{15}$
100	$5.8529 \cdot 10^{18}$

Il problema, dunque, risulta essere tanto più mal condizionato quanto più la discretizzazione è fine.

2.3 Deconvoluzione in una dimensione

In precedenza si è fatto riferimento ai problemi di deconvoluzione definendoli attraverso delle particolari equazioni lineari di Fredholm del primo tipo il cui kernel è convolutorio, cioè della forma

$$K(s, t) = h(s - t).$$

Similmente a quanto dedotto nell'esempio con cui si è concluso il paragrafo precedente, la matrice dei coefficienti derivante dalla discretizzazione attraverso l'applicazione di una regola di quadratura ha entrate del tipo

$$a_{ij} = w_j K(s_i, t_j) = w_j h(s_i - t_j).$$

Per essa, esiste dunque la fattorizzazione

$$A = HW, \quad (2.7)$$

dove

$$H = \begin{bmatrix} h(s_1 - t_1) & h(s_1 - t_2) & \dots & h(s_1 - t_n) \\ h(s_2 - t_1) & h(s_2 - t_2) & \dots & h(s_2 - t_n) \\ \vdots & \vdots & & \vdots \\ h(s_n - t_1) & h(s_n - t_2) & \dots & h(s_n - t_n) \end{bmatrix} \text{ e } W = \begin{bmatrix} w_1 & & & \\ & w_2 & & \\ & & \ddots & \\ & & & w_n \end{bmatrix}$$

rappresentano rispettivamente la matrice che discretizza il kernel e la matrice dei pesi associati al metodo di quadratura scelto.

Assumendo che i punti s_i e t_j , per $i, j = 1, \dots, n$, siano identici ed equidistanziati, cioè $s_i = t_i = \alpha + \beta i$, allora gli elementi della matrice H devono soddisfare la seguente proprietà:

$$h_{ij} = h(s_i - t_j) = h(\alpha + \beta i - \alpha - \beta j) = h(\beta i - \beta j) = h_{i-l, j-l}$$

per i, j e l opportuni.

Una matrice soddisfacente questa proprietà è una *matrice di Toeplitz*. Questa famiglia di matrici gode di importanti proprietà, le principali delle quali verranno approfondite nel prossimo paragrafo.

2.3.1 Matrici di Toeplitz e loro SVD

Definizione. Una matrice di Toeplitz è una matrice quadrata i cui elementi dipendono solamente dalla differenza tra l'indice di riga e l'indice di colonna. Sono, cioè, matrici della forma

$$T = \begin{bmatrix} t_0 & t_{-1} & t_{-2} & \dots & t_{1-n} \\ t_1 & t_0 & t_{-1} & \dots & t_{2-n} \\ t_2 & t_1 & t_0 & \dots & t_{3-n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ t_{n-1} & t_{n-2} & t_{n-3} & \dots & t_0 \end{bmatrix}.$$

Osservazione.

1. Le matrici di Toeplitz sono caratterizzate da al più $2n - 1$ elementi distinti;
2. Per opportuni indici i, j e l vale che $t_{ij} = t_{i-l, j-l} = t_{i-j}$;
3. La matrice H nella (2.7), nelle ipotesi in cui i punti s_i e t_j siano identici ed equispaziati, è una matrice di Toeplitz;

4. Le matrici di Toeplitz sono *persimmetriche*, cioè simmetriche rispetto all'anti-diagonale: $t_{ij} = t_{n-j+1, n-i+1}$.

Proprietà. L'inversa di una matrice di toeplitz T è una matrice persimmetrica.

Dimostrazione. Esprimiamo la proprietà di persimmetria della matrice T in una forma equivalente.

Consideriamo la matrice di permutazione

$$J = \begin{bmatrix} & & & & 1 \\ & & & & \\ & & & 1 & \\ & & \ddots & & \\ & & & & \\ 1 & & & & \end{bmatrix},$$

che, se moltiplicata a sinistra per T , ne inverte l'ordine delle colonne. Infatti

$$[TJ]_{ij} = \sum_{k=1}^n [T]_{ik}[J]_{kj} = [T]_{i, n+1-j} = t_{i-(n+1-j)} = t_{(i+j)-(n+1)}.$$

Osserviamo inoltre che

$$[(TJ)^t]_{ij} = [J^t T^t]_{ij} = [JT^t]_{ij} = \sum_{k=1}^n [J]_{ik}[T^t]_{kj} = \sum_{k=1}^n [J]_{ik}[T]_{jk} = T_{j, n+1-i} = t_{(i+j)-(n+1)}.$$

Pertanto

$$TJ = JT^t$$

e, poichè la matrice J è ortogonale, la proprietà di persimmetria può essere espressa dalla relazione seguente:

$$T = JT^t J. \quad (2.8)$$

Usando la (2.8) per l'inversa T^{-1} si deduce che:

$$T^{-1} = (JT^t J)^{-1} = J^{-1}(T^t)^{-1}J^{-1} = J(T^{-1})^t J,$$

ovvero che la matrice T^{-1} è persimmetrica. □

La proprietà di persimmetria di una matrice di Toeplitz può essere utilizzata per dedurre alcune relazioni di simmetria tra i vettori singolari di T .

Proprietà. A meno di un eventuale cambio di segno, gli autovettori destri di T sono uguali agli autovettori sinistri, ma con gli elementi in ordine inverso.

Dimostrazione. Consideriamo la fattorizzazione spettrale della matrice simmetrica TJ :

$$TJ = W\Lambda W^t = \sum_{j=1}^n \underline{w}_j \lambda_j \underline{w}_j^t = \sum_{j=1}^n \underline{w}_j \gamma_j \lambda_j (\gamma_j \underline{w}_j)^t, \quad (2.9)$$

dove le quantità γ_j sono scelte uguali a 1 o a -1 per rendere la quantità $\gamma_j \lambda_j$ positiva. Dalla (2.9) segue che

$$T = \sum_{j=1}^n \underline{w}_j \gamma_j \lambda_j (\gamma_j \underline{w}_j)^t J^t = \sum_{j=1}^n \underline{w}_j \gamma_j \lambda_j (\gamma_j J \underline{w}_j)^t. \quad (2.10)$$

D'altra parte, per la SVD, esistono U e V ortogonali tali che:

$$T = U \Sigma V^t = \sum_{j=1}^n \underline{u}_j \sigma_j \underline{v}_j^t. \quad (2.11)$$

Confrontando le ultime due relazioni si deducono le seguenti espressioni per i vettori singolari sinistri e destri di T :

$$\begin{cases} \underline{u}_j = \underline{w}_j \\ \underline{v}_j = \gamma_j J \underline{w}_j = \gamma_j J \underline{u}_j \end{cases} \quad (2.12)$$

per $i, j = 1, \dots, n$.

Esplicitando la seconda delle (2.12) si ottiene:

$$\begin{bmatrix} v_{1j} \\ v_{2j} \\ \vdots \\ v_{nj} \end{bmatrix} = \gamma_j \begin{bmatrix} & & & 1 \\ & & & \\ & & 1 & \\ & & \vdots & \\ 1 & & & \end{bmatrix} \begin{bmatrix} u_{1j} \\ u_{2j} \\ \vdots \\ u_{nj} \end{bmatrix} = \gamma_j \begin{bmatrix} u_{nj} \\ u_{n-1,j} \\ \vdots \\ u_{1j} \end{bmatrix}$$

cioè che

$$v_{ij} = \gamma_j u_{n-i+1,j}, \quad i, j = 1, \dots, n. \quad (2.13)$$

□

Da questa proprietà segue la seguente:

Proprietà. Se la matrice di Toeplitz T è simmetrica allora i vettori singolari destri e sinistri sono uguali eccetto che per un eventuale cambio di segno. Inoltre la sequenza degli elementi in ogni vettore è simmetrica, a meno di un cambio di segno, rispetto agli elementi centrali.

Dimostrazione. Osserviamo, innanzitutto, che per ogni matrice A simmetrica sussistono le seguenti condizioni per i suoi vettori e valori singolari:

$$(u_i, \sigma_i, v_i) = \begin{cases} (w_i, \lambda_i, w_i), & \lambda_i \geq 0 \\ (w_i, -\lambda_i, -w_i), & \lambda_i < 0. \end{cases} \quad (2.14)$$

Per (2.13) si ha inoltre che

$$v_{ij} \gamma_j u_{n-i+1,j}, \quad i, j = 1, \dots, n.$$

Allora, usando la (2.14) si deduce che

$$u_j = \tilde{\gamma}_j v_j, \quad \tilde{\gamma}_j = \pm 1$$

e quindi che

$$u_{ij} = \tilde{\gamma}_j v_{ij} = \tilde{\gamma}_j \gamma_j u_{n-i+1,j}, \quad i, j = 1, \dots, n,$$

da cui la tesi. \square

2.3.2 Matrici circolanti e convoluzione

Le matrici circolanti costituiscono una classe di matrici di Toeplitz in cui gli elementi di una colonna si ripetono nella successiva in modo “circolare”; una matrice circolante C di ordine $n \times n$ ha cioè la struttura seguente:

$$C = \begin{bmatrix} c_0 & c_{n-1} & c_{n-2} & \dots & c_1 \\ c_1 & c_0 & c_{n-1} & \dots & c_2 \\ c_2 & c_1 & c_0 & \dots & c_3 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ c_{n-1} & c_{n-2} & c_{n-3} & \dots & c_0 \end{bmatrix},$$

e risulta dunque costituita al più da n elementi distinti che soddisfano la relazione

$$c_{ij} = c_{(i-j) \bmod n}.$$

Osservazione. Le matrici circolanti possono essere usate per esprimere la convoluzione tra due segnali n -periodici, ovvero segnali i cui valori si ripetono con ciclicità ogni n intervalli spaziali o temporali.

Infatti, siano f e h due segnali n -periodici e sia g il segnale n -periodico ottenuto dalla loro convoluzione discreta, avente componenti:

$$g_i = \sum_{j=0}^{n-1} f_j h_{i-j}, \quad i = 0, \dots, n-1.$$

Se h è la prima colonna di una matrice circolante C_h allora si può scrivere che:

$$\begin{bmatrix} g_0 \\ g_1 \\ g_2 \\ \vdots \\ g_{n-1} \end{bmatrix} = \begin{bmatrix} h_0 & h_{n-1} & h_{n-2} & \dots & h_1 \\ h_1 & h_0 & h_{n-1} & \dots & h_2 \\ h_2 & h_1 & h_0 & \dots & h_3 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ h_{n-1} & h_{n-2} & h_{n-3} & \dots & h_0 \end{bmatrix} \cdot \begin{bmatrix} f_0 \\ f_1 \\ f_2 \\ \vdots \\ f_{n-1} \end{bmatrix},$$

cioè

$$g = C_h f. \tag{2.15}$$

Proposizione. Sia C_h la matrice circolante avente prima colonna $h = (h_0, h_1, \dots, h_{n-1})^t$. Allora il suo spettro, $\sigma(C_h)$, è dato dalle componenti del vettore $\mathcal{F}h$, dove \mathcal{F} indica la matrice di Fourier (1.3).

Dimostrazione. Ricordiamo innanzitutto la definizione di *prodotto di Hadamard-Schur* tra matrici, che verrà utilizzato nel seguito della dimostrazione.

Definizione. Siano

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix} \quad B = \begin{bmatrix} b_{11} & b_{12} & \dots & b_{1n} \\ b_{21} & b_{22} & \dots & b_{2n} \\ \vdots & \vdots & & \vdots \\ b_{m1} & b_{m2} & \dots & b_{mn} \end{bmatrix}$$

due matrici di ordine $m \times n$. Definiamo la matrice $A.*B$, *prodotto di Hadamard-Schur di A e B*, come:

$$A.*B = \begin{bmatrix} a_{11}b_{11} & a_{12}b_{12} & \dots & a_{1n}b_{1n} \\ a_{21}b_{21} & a_{22}b_{22} & \dots & a_{2n}b_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1}b_{m1} & a_{m2}b_{m2} & \dots & a_{mn}b_{mn} \end{bmatrix}$$

Consideriamo, allora, il segnale n -periodico ottenuto dalla convoluzione discreta dei segnali f e h e calcoliamo le entrate della sua *DFT*:

$$DFT\{g\}_i = DFT\{f * h\}_i, \quad i = 0, \dots, n-1.$$

Il primo membro della precedente relazione, per la (1.1), è identicamente uguale a:

$$DFT\{g\}_i = \hat{g}_i = \frac{1}{n} \sum_{k=0}^{n-1} g_k \omega_n^{-ik},$$

mentre il secondo membro, per il *teorema sulla DFT della convoluzione discreta* (1.9), è dato da:

$$DFT\{f * h\}_i = n DFT\{f\}_i DFT\{h\}_i = n \hat{f}_i \hat{h}_i.$$

Confrontando le due espressioni precedenti si trova allora che

$$\hat{g}_i = n \hat{f}_i \hat{h}_i, \quad i = 0, \dots, n-1,$$

ovvero, utilizzando la IDFT e le forme matriciali espresse dalle (1.5) e (1.6),

$$g = \mathcal{F}^*(n \hat{f} .* \hat{h}) = \mathcal{F}^*\left(n \cdot \frac{1}{n} \mathcal{F}h .* \frac{1}{n} \mathcal{F}f\right) = \mathcal{F}^*\left(\text{diag}(\mathcal{F}h) \cdot \frac{1}{n} \mathcal{F}f\right) = \frac{1}{n} \mathcal{F}^*\left(\text{diag}(\mathcal{F}h) \cdot \mathcal{F}f\right).$$

Pertanto

$$g = \mathcal{F}^{-1} \text{diag}(\mathcal{F}h)\mathcal{F}f. \quad (2.16)$$

Confrontando la (2.16) con la (2.15) si deduce, dunque, che

$$C_h = \mathcal{F}^{-1} \text{diag}(\mathcal{F}h)\mathcal{F}.$$

Da quest'ultima uguaglianza segue che C_h è simile a $\text{diag}(\mathcal{F}h)$ e quindi la tesi. \square

2.3.3 Moltiplicazione matrice-vettore attraverso la FFT

Il prodotto tra una matrice di Toeplitz e un vettore può essere effettuato con un costo computazionale di $O(n \log_2 n)$ flops, rispetto ai $2n^2$ che occorrono per il prodotto tra una matrice generica e un vettore.

L'idea alla base per avere prova di questa proprietà consiste nell'inserire la matrice di Toeplitz T , di ordine $n \times n$, all'interno di una matrice circolante C di ordine superiore $p \times p$, e usare l'algoritmo della FFT per implementare in modo rapido la moltiplicazione matrice-vettore con la matrice circolante C . Nello specifico, se

$$T = \begin{bmatrix} t_0 & t_{-1} & t_{-2} & \dots & t_{1-n} \\ t_1 & t_0 & t_{-1} & \dots & t_{2-n} \\ t_2 & t_1 & t_0 & \dots & t_{3-n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ t_{n-1} & t_{n-2} & t_{n-3} & \dots & t_0 \end{bmatrix},$$

allora la prima colonna di C , cioè il vettore che genera la matrice circolante, viene costruita nel modo seguente:

$$C(:, 1) = (t_0, t_1, \dots, t_{n-1}, 0, \dots, 0, t_{1-n}, \dots, t_{-1}) \in \mathbb{R}^p,$$

in cui la parte centrale è occupata da $p - 2n + 1$ zeri, essendo p la più piccola potenza di due tale che $p \geq 2n$. Il motivo per cui vengono scelte potenze di due è legato alle proprietà degli algoritmi per la FFT, i quali vengono implementati preferibilmente per potenze di 2 (si veda la descrizione dell'*algoritmo di Cooley-Tukey* nel paragrafo 1.1.5).

Pertanto, per moltiplicare T per un vettore \mathbf{x} è necessario considerare il vettore

$$\hat{\mathbf{x}} = \begin{bmatrix} \mathbf{x} \\ 0 \end{bmatrix}$$

esteso con $p - n$ zeri.

In questo modo, in virtù della presenza degli zeri e del modo con cui è stata costruita la matrice circolante C , vale che

$$C\hat{\mathbf{x}} = \begin{bmatrix} T\mathbf{x} \\ \mathbf{b} \end{bmatrix}.$$

2.4 Deconvoluzione in due dimensioni

Un problema di deconvoluzione in due dimensioni viene formulato da un'equazione integrale di Fredholm del primo tipo della forma seguente

$$\int_0^1 \int_0^1 K(x, y, x', y') f(x', y') dx' dy' = g(x, y). \quad (2.17)$$

Considerando il caso in cui il kernel K sia un operatore convolutorio reale a variabili separabili, cioè tale per cui

$$K(x, y, x', y') = h(x - x', y - y') = \kappa(x - x')\omega(y - y'),$$

essendo κ e ω due funzioni a valori reali, la (2.17) può essere riscritta in questo modo:

$$g(x, y) = \int_0^1 \kappa(x - x') \left(\int_0^1 \omega(y - y') f(x', y') dy' \right) dx'. \quad (2.18)$$

La discretizzazione di un problema di deconvoluzione in due dimensioni a variabili separabili può essere agevolmente formulata utilizzando l'operazione tra matrici del *prodotto di Kronecker* e la *vec notation*, che verranno introdotte nel paragrafo seguente.

2.4.1 Il prodotto di Kronecker e la “vec notation”

Definizione. Siano A, B due matrici rispettivamente di dimensioni $m \times n$ e $p \times q$. Il *prodotto di Kronecker* tra A e B è la matrice di dimensioni $mp \times nq$ così costruita:

$$A \otimes B = \begin{bmatrix} a_{11}B & a_{12}B & \dots & a_{1n}B \\ a_{21}B & a_{22}B & \dots & a_{2n}B \\ \vdots & \vdots & & \vdots \\ a_{m1}B & a_{m2}B & \dots & a_{mn}B \end{bmatrix},$$

o, in termini delle colonne delle due matrici $A = [a_1, a_2, \dots, a_n]$ e $B = [b_1, b_2, \dots, b_q]$,

$$A \otimes B = (a_1 \otimes b_1, \dots, a_1 \otimes b_q, a_2 \otimes b_1, \dots, a_2 \otimes b_q, \dots, a_n \otimes b_1, \dots, a_n \otimes b_q).$$

Proprietà.

1. $(A \otimes B)^t = A^t \otimes B^t$
2. Se A, B, C e D sono tali che esiste il prodotto righe per colonne tra A e C e B e D rispettivamente allora

$$(A \otimes B)(C \otimes D) = (AC) \otimes (BD)$$

Dimostrazione.

$$1. (A \otimes B)^t = \begin{bmatrix} a_{11}B & a_{12}B & \dots & a_{1n}B \\ a_{21}B & a_{22}B & \dots & a_{2n}B \\ \vdots & \vdots & & \vdots \\ a_{m1}B & a_{m2}B & \dots & a_{mn}B \end{bmatrix}^t = \begin{bmatrix} a_{11}B^t & a_{21}B^t & \dots & a_{m1}B^t \\ a_{12}B^t & a_{22}B^t & \dots & a_{m2}B^t \\ \vdots & \vdots & & \vdots \\ a_{1n}B^t & a_{2n}B^t & \dots & a_{mn}B^t \end{bmatrix} = A^t \otimes B^t$$

$$2. (A \otimes B)(C \otimes D) = \begin{bmatrix} a_{11}B & a_{12}B & \dots & a_{1n}B \\ a_{21}B & a_{22}B & \dots & a_{2n}B \\ \vdots & \vdots & & \vdots \\ a_{m1}B & a_{m2}B & \dots & a_{mn}B \end{bmatrix} \begin{bmatrix} c_{11}D & c_{12}D & \dots & c_{1s}D \\ c_{21}D & c_{22}D & \dots & c_{2s}D \\ \vdots & \vdots & & \vdots \\ c_{n1}D & c_{n2}D & \dots & c_{ns}D \end{bmatrix} =$$

$$= \begin{bmatrix} \sum_{k=1}^n a_{1k}c_{k1}BD & \sum_{k=1}^n a_{1k}c_{k2}BD & \dots & \sum_{k=1}^n a_{1k}c_{ks}BD \\ \sum_{k=1}^n a_{2k}c_{k1}BD & \sum_{k=1}^n a_{2k}c_{k2}BD & \dots & \sum_{k=1}^n a_{2k}c_{ks}BD \\ \vdots & \vdots & & \vdots \\ \sum_{k=1}^n a_{mk}c_{k1}BD & \sum_{k=1}^n a_{mk}c_{k2}BD & \dots & \sum_{k=1}^n a_{mk}c_{ks}BD \end{bmatrix} = (AC) \otimes (BD)$$

□

Osservazione. Se A e B sono matrici di Toeplitz allora la matrice $A \otimes B$ è una matrice di Toeplitz a blocchi con blocchi di Toeplitz (BTTB).

Definizione. Sia X una matrice di dimensioni $m \times n$ con colonne $X = (x_1, \dots, x_n)$. Definiamo il vettore $\text{vec}(X)$ di lunghezza mn come

$$\text{vec}(X) = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}.$$

È possibile dimostrare che il prodotto di Kronecker e la “vec” notation sono legati dalla seguente relazione:

$$(A \otimes B)\text{vec}(X) = \text{vec}(BXA^t) \quad (2.19)$$

Proposizione. Siano A e B due matrici quadrate, e siano

$$A = U\Sigma V^t, \quad B = \tilde{U}\tilde{\Sigma}\tilde{V}^t$$

le rispettive decomposizioni a valori singolari. Allora la decomposizione a valori singolari di $A \otimes B$ è

$$A \otimes B = (U \otimes \tilde{U})(\Sigma \otimes \tilde{\Sigma})(V \otimes \tilde{V})^t. \quad (2.20)$$

Dimostrazione. Utilizzando le proprietà del prodotto di Kronecker e la decomposizione a valori singolari delle matrici A e B si ha che:

$$A \otimes B = (U\Sigma V^t) \otimes (\tilde{U}\tilde{\Sigma}\tilde{V}^t) = (U\Sigma \otimes \tilde{U}\tilde{\Sigma})(V^t \otimes \tilde{V}^t) = (U \otimes \tilde{U})(\Sigma \otimes \tilde{\Sigma})(V \otimes \tilde{V}^t),$$

che rappresenta la fattorizzazione a valori singolari della matrice $A \otimes B$.

□

2.4.2 Discretizzazione per kernel con variabili separabili

Consideriamo un problema di deconvoluzione in due dimensioni con kernel convolutorio a variabili separabili reali, cioè

$$g(x, y) = \int_0^1 \kappa(x - x') \left(\int_0^1 \omega(y - y') f(x', y') dy' \right) dx'.$$

Utilizziamo, per semplicità, la regola di quadratura del punto medio per approssimare l'integrale interno. Denotando con \tilde{f} la perturbazione di f dovuta agli errori di approssimazione, si ottiene che:

$$\int_0^1 \omega(y - y') f(x', y') dy' \approx \frac{1}{n} \sum_{l=1}^n \omega(y - y'_l) \tilde{f}(x', y'_l) = \phi(x', y),$$

dove y'_l , $l = 1, \dots, n$ indicano i nodi della formula di quadratura.

Riapplicando la stessa regola all'altro integrale segue che

$$\int_0^1 \kappa(x - x') \phi(x', y) dx' \approx \frac{1}{n} \sum_{k=1}^n \kappa(x - x'_k) \phi(x'_k, y) = \psi(x, y),$$

con x'_k , $k = 1, \dots, n$ nodi della formula.

Allora, imponendo n^2 condizioni di interpolazione

$$g(x_i, y_j) = \psi(x_i, y_j) \quad \text{per } i, j = 1, \dots, n, \quad (2.21)$$

è possibile determinare le equazioni del sistema lineare corrispondente al problema di deconvoluzione considerato. Con questa scelta, il numero dei nodi della formula di quadratura e dei punti di interpolazione è lo stesso.

Siano, dunque, A, \bar{A}, F, G rispettivamente le matrici con entrate:

1. $[A]_{ik} = \frac{1}{n} \kappa(x_i - x'_k)$;
2. $[\bar{A}]_{jl} = \frac{1}{n} \omega(y_j - y'_l)$;
3. $F_{kl} = \tilde{f}(x'_k, y'_l)$;
4. $G_{ij} = g(x_i, y_j)$,

per $i, j, k, l = 1, \dots, n$. Più precisamente:

- A rappresenta il campionamento della funzione κ ;
- \bar{A} rappresenta il campionamento della funzione ω ;
- F rappresenta il campionamento della soluzione approssimata \tilde{f} ;
- G rappresenta il campionamento di g ,

e sono così strutturate:

$$A = \frac{1}{n} \begin{bmatrix} \kappa(x_1 - x'_1) & \kappa(x_1 - x'_2) & \dots & \kappa(x_1 - x'_n) \\ \kappa(x_2 - x'_1) & \kappa(x_2 - x'_2) & \dots & \kappa(x_2 - x'_n) \\ \vdots & \vdots & & \vdots \\ \kappa(x_n - x'_1) & \kappa(x_n - x'_2) & \dots & \kappa(x_n - x'_n) \end{bmatrix},$$

$$\bar{A} = \frac{1}{n} \begin{bmatrix} \omega(y_1 - y'_1) & \omega(y_1 - y'_2) & \dots & \omega(y_1 - y'_n) \\ \omega(y_2 - y'_1) & \omega(y_2 - y'_2) & \dots & \omega(y_2 - y'_n) \\ \vdots & \vdots & & \vdots \\ \omega(y_n - y'_1) & \omega(y_n - y'_2) & \dots & \omega(y_n - y'_n) \end{bmatrix},$$

$$F = \begin{bmatrix} \tilde{f}(x'_1, y'_1) & \tilde{f}(x'_2, y'_2) & \dots & \tilde{f}(x'_n, y'_n) \\ \tilde{f}(x'_2, y'_1) & \tilde{f}(x'_2, y'_2) & \dots & \tilde{f}(x'_2, y'_n) \\ \vdots & \vdots & & \vdots \\ \tilde{f}(x'_n, y'_1) & \tilde{f}(x'_n, y'_2) & \dots & \tilde{f}(x'_n, y'_n) \end{bmatrix},$$

$$G = \begin{bmatrix} g(x_1, y_1) & g(x_1, y_2) & \dots & g(x_1, y_n) \\ g(x_2, y_1) & g(x_2, y_2) & \dots & g(x_2, y_n) \\ \vdots & \vdots & & \vdots \\ g(x_n, y_1) & g(x_n, y_2) & \dots & g(x_n, y_n) \end{bmatrix}.$$

A questo punto è possibile definire una matrice $n \times n$, Φ , corrispondente all'integrazione interna. Tale matrice ha entrate

$$\Phi_{kj} = \phi(x'_k, y_j) = \frac{1}{n} \sum_{l=1}^n \omega(y_j - y'_l) \tilde{f}(x'_k, y'_l), \quad j, k = 1, \dots, n, .$$

e, pertanto,

$$\Phi = F\bar{A}^t.$$

Ragionando similmente per l'integrazione più esterna è possibile definire una matrice Ψ di ordine $n \times n$ con entrate

$$\Psi_{ij} = \psi(x_i, y_j) = \frac{1}{n} \sum_{k=1}^n \kappa(x_i - x'_k) \phi(x'_k, y_j),$$

ovvero

$$\Psi = A\Phi = AF\bar{A}^t.$$

Per la condizione (2.21) deve inoltre valere che

$$\Psi_{ij} = G_{ij}$$

per ogni $i, j = 1, \dots, n$. Segue, dunque, che

$$G = AF\bar{A}^t \Leftrightarrow \text{vec}(G) = \text{vec}(AF\bar{A}^t). \quad (2.22)$$

Usando, allora, la relazione (2.19) si deduce che

$$\text{vec}(G) = (\bar{A} \otimes A)\text{vec}(F), \quad (2.23)$$

che costituisce *il sistema di n^2 equazioni lineari associato al problema di deconvoluzione bidimensionale* (2.18).

Capitolo 3

Image restoration

Un'immagine è un segnale bidimensionale che trasporta l'informazione di un oggetto fisico e, in generale, fornisce una rappresentazione degradata di tale oggetto. La degradazione è dovuta essenzialmente a due fattori: al processo di formazione dell'immagine e al suo processo di registrazione.

La degradazione dovuta al processo di formazione dell'immagine è spesso associata al fenomeno del *blurring* (*sfocatura*), mentre la degradazione dovuta al processo di registrazione dell'immagine è associata al *rumore* ed è causata, tra i vari fattori, da errori di misura.

Il primo dei due processi è un processo deterministico, al quale può essere associato un modello matematico accurato per la sua descrizione; il secondo, al contrario, è un evento aleatorio, per il quale spesso devono essere fatte delle assunzioni in merito alle sue principali proprietà statistiche. La digital image restoration, cioè il procedimento attraverso il quale un'immagine sfocata affetta da rumore viene ripristinata sulle basi di un modello matematico che descrive il processo di sfocatura, costituisce uno degli esempi più noti di problema di deconvoluzione in due dimensioni.

La struttura della matrice che descrive il fenomeno della sfocatura dipende da quelle che vengono chiamate *boundary conditions*, cioè delle assunzioni che vengono effettuate in merito al comportamento dell'immagine nitida oltre i margini della figura. Verranno analizzati i casi corrispondenti a condizioni di Dirichlet, periodiche e riflesse.

3.1 Deduzione dell'equazione integrale di Fredholm associata al deblurring

Consideriamo la distribuzione di luce emessa da un oggetto fisico collocato in uno spazio piano, normalmente rappresentata da una funzione a valori reali in due

variabili,

$$f^{(0)}(\mathbf{x}) ,$$

dove il vettore \mathbf{x} indica la posizione dell'oggetto nel piano suddetto, cioè $\mathbf{x} = (x_1, x_2)$.

Tale oggetto si comporta rispetto al sistema di immagine come una sorgente luminosa: esso, infatti produce nel dominio dell'immagine del sistema ottico, normalmente un piano, una certa distribuzione di luce, che indichiamo con

$$g^{(0)}(\mathbf{x}).$$

Una proprietà fondamentale degli strumenti ottici è che l'immagine prodotta da una sorgente puntiforme non è un punto, ma una sorta di goccia che si irradia con una certa distribuzione.

Supponiamo che la sorgente puntiforme sia situata in un dato punto \mathbf{x} nel dominio dell'oggetto e che la sua intensità luminosa in un punto $\mathbf{x}' \neq \mathbf{x}$ venga descritta dalla distribuzione *delta di Dirac* centrata nell'origine:

$$f^{(0)}(\mathbf{x}') = \delta(\mathbf{x}' - \mathbf{x}).$$

La sua immagine, allora, sarà una funzione di \mathbf{x}' , che indichiamo con $K(\mathbf{x}, \mathbf{x}')$: essa, ragionevolmente, presenta un picco nel punto $\mathbf{x}' = \mathbf{x}$ che è circondato da una sorta di nube, dotata di una data distribuzione.

La funzione K è nota con il termine di *Point Spread Function (PSF)* (*Funzione di Diffusione del Punto*), e rappresenta, dunque, il modo con cui una sorgente puntiforme luminosa viene rilevata dal sistema ottico di acquisizione.

Poichè la PSF indica la risposta alla distribuzione delta di Dirac, essa viene anche chiamata *risposta impulsiva* del sistema. Il suo effetto è quello che propriamente descrive il fenomeno del *blurring* (sfocatura).

Consideriamo, a questo punto, un operatore A , il quale trasforma la distribuzione $f^{(0)}(\mathbf{x})$ nell'immagine sfocata $g^{(0)}(\mathbf{x})$.

Pertanto, esprimendo la distribuzione di luce emessa dall'oggetto come la sovrapposizione lineare di sorgenti luminose, cioè

$$f^{(0)}(\mathbf{x}) = \int \delta(\mathbf{x} - \mathbf{x}') f^{(0)}(\mathbf{x}') d\mathbf{x}' , \quad (3.1)$$

segue che l'immagine sfocata di tale oggetto è descritta dalla seguente equazione integrale:

$$Af^{(0)}(\mathbf{x}) = g^{(0)}(\mathbf{x}) = \int K(\mathbf{x}, \mathbf{x}') f^{(0)}(\mathbf{x}') d\mathbf{x}' , \quad (3.2)$$

che ha la forma di un' *equazione integrale di Fredholm del primo tipo in due variabili*.

A questo punto, assumendo che il sistema di elaborazione di immagini sia *invariante per traslazioni*, cioè a dire la PSF che appare nella (3.2) soddisfa la relazione

$$K(\mathbf{x}, \mathbf{x}_0) = K(\mathbf{x} - \mathbf{x}_0) ,$$

la (3.2) assume la forma di un'equazione integrale di Fredholm del primo tipo con kernel convolutorio:

$$g^{(0)}(\mathbf{x}) = \int K(\mathbf{x} - \mathbf{x}')f^{(0)}(\mathbf{x}')d\mathbf{x}' . \quad (3.3)$$

L'immagine sfocata è quindi ottenuta dal *prodotto convolutorio* tra la PSF e la distribuzione $f^{(0)}$, ovvero

$$g^{(0)} = K * f^{(0)} .$$

Per risalire alla $f^{(0)}(\mathbf{x}')$, note $K(\mathbf{x} - \mathbf{x}')$ e $g^{(0)}(\mathbf{x})$, dovrà quindi essere risolto un *problema di deconvoluzione* della forma (2.4).

Per questo motivo, il fenomeno del deblurring viene spesso associato al problema dell'*image deconvolution*.

3.2 Il problema dell'immagine deconvolution

Ricordiamo alcuni notevoli risultati sulla teoria delle *trasformate di Fourier*, che verranno utilizzati nei paragrafi che seguono.

Teorema. (*Trasformata di Fourier della convoluzione*)

Siano f e g due funzioni nello spazio L^1 e sia $f * g$ la loro convoluzione. Allora allora vale che

$$\mathcal{F}\{f * g\} = \mathcal{F}\{f\} \cdot \mathcal{F}\{g\} = \hat{f}\hat{g} .$$

Definizione. Una funzione $f: \mathbb{R}^n \rightarrow \mathbb{C}^n$ si dice *anti-trasformabile secondo Fourier* se esiste finito per ogni $\mathbf{x} \in \mathbb{R}^n$ il seguente integrale in valore principale:

$$\frac{V.P.}{(2\pi)^n} \int_{\mathbb{R}^n} f(\boldsymbol{\omega})e^{i\boldsymbol{\omega} \cdot \mathbf{x}}d\boldsymbol{\omega} .$$

Teorema. (*Uguaglianza di Parseval generalizzata*)

Siano $f, h \in L^1(\mathbb{R}^n)$. Allora

$$\int_{\mathbb{R}^n} f(\mathbf{x})h^*(\mathbf{x})d\mathbf{x} = \frac{1}{(2\pi)^n} \int_{\mathbb{R}^n} \hat{f}(\boldsymbol{\omega})\hat{h}(\boldsymbol{\omega})d\boldsymbol{\omega} ,$$

dove h^* indica il coniugato di h .

Nel paragrafo che segue ammettiamo che la validità del concetto di trasformata e i risultati ricordati siano estesi per funzioni nello spazio di Hilbert L^2 , fatto che può essere rigorosamente giustificato lavorando nello *spazio delle funzioni a decrescenza rapida*.

3.2.1 Formulazione del problema nel continuo

Consideriamo un sistema di immagine a invarianza spaziale, modellizzato da un'equazione della forma (3.3).

Date f, K funzioni appartenenti allo spazio di Hilbert $L^2(\mathbb{R}^2)$, definiamo l'operatore di convoluzione

$$A: L^2(\mathbb{R}^2) \rightarrow L^2(\mathbb{R}^2), \quad f \mapsto Af$$

tale che

$$(Af)(\mathbf{x}) = \iint_{\mathbb{R}^2} K(\mathbf{x} - \mathbf{x}')f(\mathbf{x}')d\mathbf{x}', \quad (3.4)$$

o equivalentemente

$$(Af)(\mathbf{x}) = (K * f)(\mathbf{x}).$$

L'operatore A viene detto *operatore associato alla PSF*.

Applicando il *teorema della trasformata di Fourier della convoluzione* al secondo membro dell'ultima relazione ottenuta, si deduce che

$$(\mathcal{F}\{K * f\})(\boldsymbol{\omega}) = \mathcal{F}\{K\}(\boldsymbol{\omega}) \cdot \mathcal{F}\{f\}(\boldsymbol{\omega}) = \hat{K}(\boldsymbol{\omega})\hat{f}(\boldsymbol{\omega}).$$

Allora, sfruttando la *formula di inversione per la trasformata di Fourier* si ottiene che:

$$(Af)(\mathbf{x}) = \frac{1}{(2\pi)^2} \int_{\mathbb{R}^2} \hat{K}(\boldsymbol{\omega}) \cdot \hat{f}(\boldsymbol{\omega})e^{i\mathbf{x}\cdot\boldsymbol{\omega}}d\boldsymbol{\omega}. \quad (3.5)$$

Effettuiamo, a questo punto, alcune ipotesi sull'operatore A . Assumiamo che sia:

- *lineare*, per cui $A(\alpha_1 f^{(1)} + \alpha_2 f^{(2)}) = \alpha_1 Af^{(1)} + \alpha_2 Af^{(2)}$, $\alpha_1, \alpha_2 \in \mathbb{R}$
- *limitato*.

Pertanto, se

$$\hat{K}_{\max} = \max_{\boldsymbol{\omega}} |\hat{K}(\boldsymbol{\omega})|,$$

allora, passando alla norma-2,

$$\|(Af)_{\mathbf{x}}\| = \frac{1}{(2\pi)^2} \left(\int_{\mathbb{R}^2} |\hat{K}(\boldsymbol{\omega})\hat{f}(\boldsymbol{\omega})e^{i\mathbf{x}\cdot\boldsymbol{\omega}}|^2 d\boldsymbol{\omega} \right)^{\frac{1}{2}} \leq \frac{1}{(2\pi)^2} \hat{K}_{\max} \left(\int_{\mathbb{R}^2} |\hat{f}(\boldsymbol{\omega})e^{i\mathbf{x}\cdot\boldsymbol{\omega}}|^2 d\boldsymbol{\omega} \right)^{\frac{1}{2}},$$

cioè

$$\|(Af)_{\mathbf{x}}\| \leq \frac{1}{(2\pi)^2} \hat{K}_{\max} \left(\int_{\mathbb{R}^2} |\hat{f}(\boldsymbol{\omega})|^2 d\boldsymbol{\omega} \right)^{\frac{1}{2}}.$$

Sfruttando, infine, l'*uguaglianza di Parseval*, si deduce che

$$\|(Af)_{\mathbf{x}}\| \leq \hat{K}_{\max} \left(\int_{\mathbb{R}^2} |f(x)|^2 dx \right)^{\frac{1}{2}} = \|f\| \hat{K}_{\max}.$$

Osservazione. \hat{K}_{\max} rappresenta la *norma spettrale* dell'operatore A , cioè \hat{K}_{\max} è la più piccola costante C tale che

$$\|Af\| \leq C \|f\|.$$

Così come A è un operatore limitato associato a K , allo stesso modo il suo *operatore aggiunto* A^* è un operatore limitato associato a K^* , complesso coniugato di K .

Definizione. Si dice *operatore aggiunto* di A l'operatore A^* tale che

$$\langle Af, g \rangle = \langle f, A^*g \rangle$$

Osserviamo che:

$$\langle Af, g \rangle = \int_{\mathbb{R}^2} (Af)(\mathbf{x})g^*(\mathbf{x})d\mathbf{x} = \frac{1}{(2\pi)^2} \int_{\mathbb{R}^2} (\widehat{Af})(\boldsymbol{\omega})\hat{g}^*(\boldsymbol{\omega})d\boldsymbol{\omega},$$

per l'uguaglianza di Parseval.

Utilizzando la (3.5) e applicando il *teorema di convoluzione* si deduce che

$$\begin{aligned} \langle Af, g \rangle &= \frac{1}{(2\pi)^2} \int_{\mathbb{R}^2} (\widehat{K * f})(\boldsymbol{\omega})\hat{g}^*(\boldsymbol{\omega})d\boldsymbol{\omega} = \frac{1}{(2\pi)^2} \int_{\mathbb{R}^2} \hat{K}(\boldsymbol{\omega})\hat{f}(\boldsymbol{\omega})\hat{g}^*(\boldsymbol{\omega})d\boldsymbol{\omega} = \\ &= \frac{1}{(2\pi)^2} \int_{\mathbb{R}^2} \hat{f}(\boldsymbol{\omega})\left(\hat{K}^*(\boldsymbol{\omega})\hat{g}(\boldsymbol{\omega})\right)^* d\boldsymbol{\omega}. \end{aligned}$$

Utilizzando la formula di inversione per la trasformata di Fourier segue allora che

$$\langle Af, g \rangle = \int_{\mathbb{R}^2} f(\mathbf{x})\mathcal{F}^{-1}\{(\hat{K}(\boldsymbol{\omega})\hat{g}(\boldsymbol{\omega}))^*\}(\mathbf{x})d\mathbf{x} = \int_{\mathbb{R}^2} f(\mathbf{x})(A^*g)^*(\mathbf{x})d\mathbf{x} = \langle f, A^*g \rangle$$

Confrontando le relazioni precedenti, si deduce quindi che

$$A^*g = \mathcal{F}^{-1}\{(\hat{K}(\boldsymbol{\omega})\hat{g}(\boldsymbol{\omega}))\}(\mathbf{x})$$

la quale, a sua volta, per il *teorema di convoluzione*, può essere riscritta nel modo seguente:

$$A^*g = (K^* * g)(\mathbf{x}) = \int_{\mathbb{R}^2} K^*(\mathbf{x}' - \mathbf{x})g(\mathbf{x}')d\mathbf{x}'.$$

Quest'ultima relazione indica, dunque, che anche l'operatore aggiunto A^* è un operatore di convoluzione limitato, associato alla PSF K^* .

3.2.2 Formulazione del problema nel discreto

Analoghi risultati a quelli dedotti nel paragrafo precedente, possono essere ottenuti per immagini e oggetti discreti bidimensionali.

In questo caso l'operatore A associato alla PSF K è definito come

$$(Af)_{m,n} = \sum_{k,l=0}^{N-1} K_{m-k,n-l} f_{k,l},$$

che esprime la convoluzione discreta di K e f nel caso bidimensionale.

Lo spazio dell'oggetto e quello dell'immagine sono, in questo caso, degli spazi euclidei di vettori bidimensionali, muniti del prodotto scalare standard.

In sintesi, dunque, il problema dell'immagine deconvolution può essere formulato come il problema di stimare l'oggetto $f^{(0)}(\mathbf{x})$ (o la sua versione discreta $\mathbf{f}^{(0)}$), note l'immagine, affetta da rumore, $g(\mathbf{x})$ (o la sua versione discreta \mathbf{g}), e la PSF $K(\mathbf{x})$ (o la sua versione discreta \mathbf{K} .)

3.2.3 Esempi di blurring

I motivi per cui un'immagine può risultare sfocata possono essere diversi. In questa sezione discuteremo alcuni dei più importanti, analizzando la struttura delle corrispondenti PSF, cioè il kernel $K(x, y, x', y')$ che modella il fenomeno. Per semplicità supponiamo che $K(x, y, x', y') = \kappa(x - x')\omega(y - y')$, cioè la sfocatura, è identica in tutte le parti dell'immagine e si separa in una componente orizzontale e una verticale.

Linear motion blur

La *sfocatura da movimento rettilineo* si verifica quando tra il sistema di registrazione dell'immagine e l'oggetto da fotografare si ha un moto relativo. Nel caso in cui tale moto sia puramente orizzontale, la luce emessa da una sorgente puntiforme viene acquisita come un segmento e la PSF ha la forma

$$K(x, y, x', y') = h_L(x - x') = \begin{cases} \frac{1}{2L}, & |x - x'| \leq L \\ 0, & \text{altrimenti} \end{cases}, \quad (3.6)$$

essendo L il parametro che caratterizza la sbavatura dell'immagine.

Similmente, se il moto è verticale la PSF è data da:

$$K(x, y, x', y') = h_L(y - y') = \begin{cases} \frac{1}{2L}, & |y - y'| \leq L \\ 0, & \text{altrimenti} \end{cases}$$

Sfocatura da non messa a fuoco

Nel caso di una lente, un oggetto puntiforme è a fuoco se la sua distanza dalla lente, d_0 , soddisfa la legge dei punti coniugati

$$\frac{1}{d_0} + \frac{1}{d_i} = \frac{1}{d_f}, \quad (3.7)$$

essendo d_i la distanza tra la lente e il piano dell'immagine e d_f la distanza focale della lente. Se questa condizione è soddisfatta allora, per le leggi dell'ottica geometrica, l'immagine di un punto è ancora un punto. Al contrario, se d_0 non soddisfa la condizione (3.7) allora l'immagine di un punto è un disco, il cui raggio si dimostra poter essere calcolato in funzione dei parametri d_0 , d_i , d_f e del raggio effettivo della lente.

Un buon modello di PSF per il caso in cui la sfocatura sia causata dalla non messa a fuoco è

$$K(x, y, x', y') = \begin{cases} \frac{1}{\pi R^2}, & (x - x')^2 + (y - y')^2 \leq R^2 \\ 0, & \text{altrimenti} \end{cases}, \quad (3.8)$$

dove il parametro R caratterizza la sfocatura (cfr. [11], [14], [15]).

Osserviamo, tuttavia, che questa funzione non è separabile rispetto alle variabili $x - x'$ e $y - y'$. Per questo motivo, nella pratica, si utilizza una PSF alternativa del tipo

$$K(x, y, x', y') = h_R(x - x')h_R(y - y'),$$

dove la funzione h_R è la stessa presente nella (3.6)

Sfocatura originata da turbolenza atmosferica

Questo caso è tipico, per esempio, nell'astronomia osservazionale e nel telerilevamento, in cui l'oggetto che si intende riprendere è soggetto agli effetti della turbolenza atmosferica, la quale genera variazioni randomiche nell'indice di rifrazione del mezzo in cui si propaga il segnale luminoso, determinando un'alterazione del percorso del raggio di luce emesso dalla sorgente.

Per ragioni di natura pratica, la PSF viene modellizzata attraverso un modello Gaussiano come il seguente:

$$K(x, y, x', y') = \frac{1}{2\pi\sigma\bar{\sigma}} e^{-\frac{1}{2}\left(\frac{x-x'}{\sigma}\right)^2 - \frac{1}{2}\left(\frac{y-y'}{\bar{\sigma}}\right)^2}, \quad (3.9)$$

essendo σ e $\bar{\sigma}$ due costanti che caratterizzano, rispettivamente, la sfocatura lungo l'asse x e l'asse y .

3.3 Boundary conditions

Nei punti che si trovano nei pressi dei margini di un'immagine la discussione della PSF così come è stata fatta negli esempi di blurring presentati nel paragrafo precedente potrebbe risultare più problematica: infatti, ciò che si verifica, è che l'informazione proveniente dall'immagine nitida fuoriesca dai margini dell'immagine sfocata, determinando una perdita di informazione.

Un buon modello per l'immagine deblurring deve chiaramente tenere in considerazione questo aspetto e analizzarne gli aspetti ai margini dell'immagine.

La tecnica più frequentemente adottata per ovviare alla perdita di informazione nel contorno è quella di effettuare delle assunzioni sul comportamento dell'immagine nitida oltre i margini dell'immagine: devono cioè essere imposte quelle che vengono chiamate *boundary conditions*.

3.3.1 Boundary conditions per immagini in una dimensione

Seguendo la trattazione presente in [2], consideriamo inizialmente il problema del deblurring in una dimensione.

Siano

$$\tilde{f} = (\dots, f_{-m+1}, \dots, f_0, f_1, \dots, f_n, f_{n+1}, \dots, f_{n+m}, \dots)^t$$

il segnale completo, comprendente l'informazione proveniente sia dall'interno dell'immagine che oltre i suoi margini, e

$$h = (\dots, 0, 0, h_{-m}, h_{-m+1}, \dots, h_0, \dots, h_{m-1}, h_m, 0, 0, \dots)^t$$

la funzione che descrive il fenomeno del blurring.

Il segnale sfocato g sappiamo poter essere calcolato attraverso la convoluzione discreta di f e h .

Pertanto, se

$$f = (f_1, f_2, \dots, f_n)^t$$

costituisce il segnale da ricostruire allora il problema in forma matriciale può essere espresso come segue:

$$T_r = \begin{bmatrix} & & & & 0 \\ h_{-m} & & & & \\ \vdots & \ddots & & & \\ h_{-1} & \dots & \dots & \dots & h_{-m} \end{bmatrix}_{n \times m}, \quad f_r = \begin{bmatrix} f_{n+1} \\ f_{n+2} \\ \vdots \\ f_{n+m-1} \\ f_{n+m} \end{bmatrix}_{m \times 1}$$

Zero boundary condition

Nella *zero boundary condition*, o condizione di Dirichlet, si assume che l'immagine esatta sia nera fuori dai margini, cioè che

$$f_l = f_r = \mathbf{0}.$$

Pertanto, il sistema matriciale (3.10) assume la forma

$$Tf = g, \quad (3.12)$$

essendo la matrice dei coefficienti T una *matrice di Toeplitz*.

La zero boundary condition è una buona scelta quando l'immagine esatta è buia oltre i margini, come ad esempio accade per molte immagini astronomiche aventi sfondo nero. Negli altri casi, una sua assunzione produce effetti negativi nella ricostruzione.

Periodic boundary condition

Nella condizione di boundary condition periodica si assume che i valori del segnale al di fuori del dominio dell'immagine si ripetano, cioè

$$f_j = f_{n-j}$$

per ogni j nella (3.10).

Pertanto, il sistema matriciale (3.10) assume la forma seguente:

$$Bf = [(0|T_l) + T + (T_r|0)]f, \quad (3.13)$$

dove $(0|T_l)$ e $(T_r|0)$ sono rispettivamente le matrici di Toeplitz di ordine $n \times n$ ottenute aggiungendo $(n - m)$ colonne di zeri alle matrici T_l e T_r .

La matrice dei coefficienti B risulta essere una matrice circolante, infatti:

$$\begin{bmatrix} 0 & \dots & 0 & h_m & \dots & h_1 \\ & & & & \ddots & \vdots \\ & & & & & h_m \\ 0 & \dots & 0 & 0 & \dots & 0 \\ \vdots & & \vdots & \vdots & & \vdots \\ 0 & \dots & 0 & 0 & \dots & 0 \end{bmatrix} + \begin{bmatrix} h_0 & \dots & h_{-m} & & 0 \\ \vdots & \ddots & \ddots & \ddots & \\ h_m & \ddots & \ddots & \ddots & h_{-m} \\ & \ddots & \ddots & \ddots & \vdots \\ 0 & & h_m & \dots & h_0 \end{bmatrix} +$$

$$+ \begin{bmatrix} & & & 0 & 0 & \dots & 0 \\ & & & & & & \\ h_{-m} & & & & & & \\ \vdots & \ddots & & & & & \\ h_{-1} & \dots & h_{-m} & 0 & \dots & 0 \end{bmatrix} = \begin{bmatrix} h_0 & h_{-1} & \dots & h_{-m} & h_m & \dots & h_1 \\ h_1 & h_0 & \dots & h_{-m+1} & h_{-m} & \dots & h_2 \\ \vdots & \vdots & & \vdots & \vdots & & \vdots \\ h_m & h_{m-1} & \dots & & & & h_{-m} \\ h_{-m} & h_m & \dots & & & & h_{-m+1} \\ \vdots & \vdots & \dots & & & & \vdots \\ h_{-1} & h_{-2} & \dots & & & & h_0 \end{bmatrix}$$

Reflexive boundary condition

Nella boundary condition riflessiva, o Neumann boundary condition, si assume che i dati all'esterno dei margini dell'immagine siano una "riflessione" di quelli all'interno, cioè, si pone:

$$\begin{aligned} f_0 &= f_1 \\ f_{-1} &= f_2 \\ &\vdots \\ f_{-m+1} &= f_m \end{aligned}$$

e

$$\begin{aligned} f_{n+1} &= f_n \\ f_{n+2} &= f_{n-1} \\ &\vdots \\ f_{n+m} &= f_{n-m+1} \end{aligned}$$

Pertanto il sistema matriciale (3.10) assume la forma seguente:

$$Af = [(0|T_l)J + T + (T_r|0)J]f = g, \tag{3.14}$$

dove J è la matrice di permutazione
$$\begin{bmatrix} 0 & \dots & 0 & 1 \\ 0 & \dots & 1 & 0 \\ & & \ddots & \\ 1 & \dots & 0 & 0 \end{bmatrix}.$$

La matrice dei coefficienti che si ottiene è una matrice data dalla somma tra una matrice di Toeplitz e una matrice di Hankel, intendendo per matrice di Hankel una matrice le cui entrate sono costanti su ogni antidiagonale.

Infatti:

$$\begin{aligned}
 (0|T_l)J+T+(T_r|0)J &= \begin{bmatrix} h_1 & h_2 & \dots & h_m & 0 & \dots & 0 \\ h_2 & h_3 & & 0 & 0 & \dots & 0 \\ \vdots & \vdots & & \vdots & \vdots & & \vdots \\ h_m & 0 & \dots & 0 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 & 0 & \dots & 0 \\ \vdots & \vdots & & \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & 0 & 0 & \dots & 0 \end{bmatrix} + \begin{bmatrix} h_0 & \dots & h_{-m} & & 0 \\ \vdots & \ddots & \ddots & \ddots & \\ h_m & \ddots & \ddots & \ddots & h_{-m} \\ & \ddots & \ddots & \ddots & \vdots \\ 0 & & h_m & \dots & h_0 \end{bmatrix} + \\
 &+ \begin{bmatrix} 0 & \dots & \dots & & 0 \\ \vdots & & & & \vdots \\ 0 & \dots & \dots & & h_{-m} \\ \vdots & & & & \vdots \\ \vdots & & & & h_{-2} \\ 0 & \dots & 0 & h_{-m} & \dots & h_{-2} & h_{-1} \end{bmatrix} = \\
 &= \begin{bmatrix} h_1 & h_2 & \dots & h_m & 0 & \dots & 0 \\ h_2 & h_3 & \dots & 0 & 0 & \dots & 0 \\ \vdots & \vdots & & \vdots & \vdots & & \vdots \\ h_m & 0 & \dots & 0 & 0 & \dots & 0 \\ 0 & \dots & & & \dots & h_{-m} \\ \vdots & & & & & \vdots \\ \vdots & & & & & h_{-2} \\ 0 & \dots & 0 & h_{-m} & \dots & h_{-1} \end{bmatrix} + \begin{bmatrix} h_0 & \dots & h_{-m} & & 0 \\ \vdots & \ddots & \ddots & \ddots & \\ h_m & \ddots & \ddots & \ddots & h_{-m} \\ & \ddots & \ddots & \ddots & \vdots \\ 0 & & h_m & \dots & h_0 \end{bmatrix}
 \end{aligned}$$

dove nell'ultima somma la prima delle due matrici è una matrice di Hankel, la seconda di Toeplitz.

3.3.2 Boundary conditions per immagini bidimensionali

Utilizzando le proprietà del prodotto di Kronecker illustrate in precedenza e la relazione (2.23), i risultati ottenuti nel paragrafo precedente possono essere estesi al caso bidimensionale.

Nel dettaglio:

1. se si sceglie una *zero boundary condition*, la matrice dei coefficienti è una *matrice di Toeplitz a blocchi con blocchi di Toeplitz* (BTTB), essendo il prodotto di Kronecker tra matrici di Toeplitz una matrice avente questa struttura;

2. similmente, se si sceglie una *periodic boundary condition* la matrice dei coefficienti, data dal prodotto di Kronecker tra due matrici circolanti, è una *matrice circolante a blocchi con blocchi circolanti* (BCCB);
3. nel caso di *reflexive boundary condition* la matrice dei coefficienti, ottenuta dal prodotto di Kronecker tra due matrici della forma Toeplitz+Hankel, è data dalla somma di quattro matrici:
 - una *matrice di Toeplitz a blocchi con blocchi di Toeplitz* (BTTB);
 - una *matrice di Toeplitz a blocchi con blocchi di Hankel* (BTHB);
 - una *matrice di Hankel a blocchi con blocchi di Toeplitz* (BHTB);
 - una *matrice di Hankel a blocchi con blocchi di Hankel* (BHHB).

Una conseguenza fondamentale della struttura matriciale del problema, derivante dall'assunzione delle particolari boundary conditions, è legata alle proprietà algebriche della matrice dei coefficienti e dei relativi vantaggi computazionali che ne conseguono. In particolare, valgono i seguenti risultati.

Proposizione. Sia A una matrice BCCB di ordine $n^2 \times n^2$. Allora ammette fattorizzazione spettrale

$$A = \mathcal{F}_2^* \Lambda \mathcal{F}_2,$$

dove \mathcal{F}_2 indica la matrice associata alla trasformata di Fourier discreta bidimensionale (DFT2).

Dimostrazione. Mostriamo innanzitutto che A è normale.

Poichè A è BCCB allora esistono C_1 e C_2 matrici circolanti di ordine n tali che

$$A = C_1 \otimes C_2 = (\mathcal{F}^* \Lambda_1 \mathcal{F}) \otimes (\mathcal{F}^* \Lambda_2 \mathcal{F}), \quad (3.15)$$

dove nell'ultima uguaglianza sono state utilizzate le fattorizzazioni spettrali di C_1 e C_2 rispettivamente.

Pertanto, sfruttando le proprietà del prodotto di Kronecker, si deduce che

$$A^* = (C_1 \otimes C_2)^* = C_1^* \otimes C_2^* = (\mathcal{F}^* \Lambda_1^* \mathcal{F}) \otimes (\mathcal{F}^* \Lambda_2^* \mathcal{F}).$$

Quindi:

$$AA^* = (\mathcal{F}^* \Lambda_1 \mathcal{F} \mathcal{F}^* \Lambda_1^* \mathcal{F}) \otimes (\mathcal{F}^* \Lambda_2 \mathcal{F} \mathcal{F}^* \Lambda_2^* \mathcal{F}).$$

Poichè $\mathcal{F} \mathcal{F}^* = nI$ segue che

$$\begin{aligned} AA^* &= (n * \mathcal{F}^* \Lambda_1 \Lambda_1^* \mathcal{F}) \otimes (n * \mathcal{F}^* \Lambda_2 \Lambda_2^* \mathcal{F}) = (n * \mathcal{F}^* \Lambda_1^* \Lambda_1 \mathcal{F}) \otimes (n * \mathcal{F}^* \Lambda_2^* \Lambda_2 \mathcal{F}) = \\ &= (\mathcal{F}^* \Lambda_1^* \mathcal{F} \mathcal{F}^* \Lambda_1 \mathcal{F}) \otimes (\mathcal{F}^* \Lambda_2^* \mathcal{F} \mathcal{F}^* \Lambda_2 \mathcal{F}) = A^* A. \end{aligned}$$

Allora, A ammette fattorizzazione spettrale, che, per la (3.15), risulta essere data da

$$A = (\mathcal{F} \otimes \mathcal{F})^*(\Lambda_1 \otimes \Lambda_2)(\mathcal{F} \otimes \mathcal{F}),$$

da cui la tesi. \square

In maniera analoga, è possibile dimostrare che, sotto particolari condizioni di simmetria della PSF, una matrice A con struttura BTTB+BTHB+BHTB+BHHB ha fattorizzazione spettrale del tipo

$$A = \mathcal{C}^t \Lambda \mathcal{C},$$

con \mathcal{C} matrice associata alla trasformata discreta dei coseni bidimensionale (DCT2). La PSF di tipo gaussiano (cfr. (3.9)) rientra in questa casistica.

Osservazione. La particolare struttura matriciale della BCCB (risp. BTTB+BTHB+BHTB+BHHB con PSF gaussiana) consente di:

1. implementare i prodotti matrice-vettore abbattendo i costi computazionali grazie alla FFT (risp. DCT);
2. Calcolare gli autovalori della matrice conoscendo soltanto la prima colonna, \mathbf{a}_1 , della BCCB (risp. \mathbf{a}_1 , della BTTB+BTHB+BHTB+BHHB) e implementando il prodotto matrice-vettore $\sqrt{n}\mathcal{F}_2\mathbf{a}_1$ (risp. $\mathcal{C}\mathbf{a}_1$).

3.4 Proprietà numeriche del problema dell'immagine deconvolution

In analogia con quanto accennato nel secondo capitolo, il problema dell'immagine deconvolution, derivando da un'equazione integrale di Fredholm del primo tipo presenta le due seguenti proprietà numeriche:

- il problema, nel caso continuo, è in generale un problema *mal posto*¹;
- nel discreto, il problema, anche quando viene formulato come problema ben posto, è *mal condizionato*.

¹In accordo con la definizione classica nel senso di Hadamard, intendiamo che un problema è *ben posto* se: ammette soluzione unica; tale soluzione esiste indipendentemente dai dati e dipende con continuità da essi.

3.4.1 Mal posizione del problema nel continuo

Supponiamo di voler risolvere il problema della forma (3.4), cioè

$$Af = g,$$

con A operatore lineare e limitato.

Esaminiamo in quali casi la soluzione per tale problema è unica. Premettiamo il seguente importante risultato:

Proposizione. La soluzione del problema $Af = g$ è unica se e solamente se il corrispondente problema omogeneo, $Af = 0$, ammette come unica soluzione la soluzione banale $f = 0$, cioè a dire, l'operatore A è iniettivo.

Dimostrazione. Dimostriamo innanzitutto che se il problema omogeneo ha come unica soluzione la soluzione banale allora esiste un'unica soluzione per $Af = g$.

Supponiamo per assurdo che $Af = g$ abbia due soluzioni: f_1 e f_2 , $f_1 \neq f_2$. Allora

$$A(f_1 - f_2) = 0 \implies f = f_1 - f_2 \in \mathcal{N}(A) = \{f \in L^2(\mathbb{R}) : Af = 0\},$$

cioè il problema omogeneo associato ammetterebbe anche una soluzione non banale, ma ciò è contro l'ipotesi.

Viceversa, ragionando sempre per assurdo, supponiamo che esistano $f \in L^2(\mathbb{R}^2)$, $f \neq 0$ e $f_1 \in L^2(\mathbb{R}^2)$ tali che $Af = 0$ e $Af_1 = g$. Si consideri $f_2 = f_1 + f$. Allora, sfruttando la linearità di A ,

$$Af_2 = Af_1 + Af = Af_1 = g.$$

Pertanto la soluzione del problema $Af = g$ non sarebbe unica, ma ciò è contro l'ipotesi. □

La proposizione precedente garantisce, dunque, che $Af = g$ ha soluzione unica se e solo se

$$\mathcal{N}(A) = \{\mathbf{0}\}.$$

Tuttavia

$$\mathcal{N}(A) = \{f \in L^2(\mathbb{R}^2) : Af = 0\} = \{f \in L^2(\mathbb{R}^2) : K * f = 0\},$$

cioè, passando alla trasformata di Fourier e sfruttando il teorema di convoluzione, la condizione affinché $f \in \mathcal{N}(A)$ è equivalente a richiedere che

$$\hat{K}(\omega)\hat{f}(\omega) = 0. \tag{3.16}$$

Perchè valga la (3.16) deve dunque valere che $\forall \omega \in \Omega$, $\hat{K}(\omega) = 0$ oppure $\hat{f}(\omega) = 0$.

A questo punto è necessario distinguere due casi:

1. $\text{supp}(\hat{K}(\omega)) = \overline{\{\omega \in \Omega : \hat{K}(\omega) \neq 0\}} = \Omega$. In questo caso necessariamente deve valere che $\hat{f}(\omega) = 0$, pertanto la soluzione del problema è unica.
2. $\text{supp}(\hat{K}(\omega)) \subset \Omega$ allora $\hat{k}(\omega) = 0$ per qualche $\omega \in \Omega$. In questa circostanza $\mathcal{N}(A) \neq \{\mathbf{0}\}$, e, pertanto, la soluzione non è unica e il problema risulta mal posto.

Studiamo, ora, il problema legato all'esistenza della soluzione nel caso in cui sia soddisfatta l'unicità, cioè quando $\text{supp}(\hat{K}(\omega)) = \Omega$.

In questo caso, essendo $\hat{g}(\omega) = \hat{K}(\omega)\hat{f}(\omega)$ segue che

$$\hat{f}(\omega) = \frac{\hat{g}(\omega)}{\hat{K}(\omega)}. \quad (3.17)$$

Considerando l'immagine affetta da rumore, cioè

$$\hat{g}(\omega) = \hat{K}(\omega)\hat{f}^{(0)} + \hat{w}(\omega),$$

e sostituendo nella (3.17), si ottiene

$$\hat{f}(\omega) = \hat{f}^{(0)}(\omega) + \frac{\hat{w}(\omega)}{\hat{K}(\omega)}. \quad (3.18)$$

Allora, se:

1. $\hat{K}(\omega) = 0$ per qualche $\omega \in \Omega$, $\hat{f}(\omega)$ presenta una singolarità nei punti in cui la trasformata di Fourier di K si annulla. Pertanto, la trasformata di Fourier inversa di $\hat{f}(\omega)$ potrebbe non esistere in $L^2(\mathbb{R}^2)$, e dunque il problema $Af = g$ potrebbe non avere soluzioni.
2. $\hat{K}(\omega) \rightarrow 0$ per $|\omega| \rightarrow \infty$, ciò che normalmente si verifica è che $\nexists \lim_{|\omega| \rightarrow \infty} \frac{\hat{w}(\omega)}{\hat{K}(\omega)}$, pertanto, anche in questo caso, l'inversa della trasformata $\hat{f}(\omega)$ non esiste, e, di conseguenza, il problema $Af = g$ potrebbe non avere soluzioni.

Considerando questi risultati, si può concludere che la soluzione al problema esiste ed è una funzione appartenente allo spazio $L^2(\mathbb{R}^2)$ se e solo se l'immagine $g(x)$ soddisfa la condizione

$$\int_{\mathbb{R}^2} \left| \frac{\hat{g}(\omega)}{\hat{K}(\omega)} \right|^2 d\omega < \infty, \quad (3.19)$$

ovvero $\frac{\hat{g}(\omega)}{\hat{K}(\omega)}$ è una funzione a quadrato sommabile. Questa condizione, infatti, consente di calcolare la trasformata inversa nella (3.18) e di determinare la soluzione al problema. Tuttavia, la condizione (3.19) non sempre è soddisfatta, e dunque, in generale, anche nel caso in cui siano verificate le condizioni per l'unicità, la soluzione potrebbe non esistere e il problema sarebbe comunque mal posto.

3.4.2 Discretizzazione e mal condizionamento

Per quanto riguarda la discretizzazione di un problema mal posto è necessario precisare innanzitutto che il corrispondente problema nel discreto

$$\mathbf{A}\mathbf{f} = \mathbf{g}$$

potrebbe essere ben posto. Infatti, i punti $\omega \in \Omega$ tali che $\hat{K}(\omega) = 0$ potrebbero non appartenere alla particolare discretizzazione considerata, e, pertanto le condizioni per l'unicità della soluzione sarebbero sempre garantite.

Inoltre, in tale caso, la soluzione è sempre determinabile e dipende con continuità dai dati.

Infatti, se $\mathbf{A}\mathbf{f} = \mathbf{g}$ ammette un'unica soluzione allora $\hat{K}_{k,l} \neq 0 \forall k, l$ e, pertanto

$$\hat{f}_{k,l} = \frac{\hat{g}_{k,l}}{\hat{K}_{k,l}}$$

risulta ben definita.

Allora, usando la *IDFT bidimensionale*, si deduce che la soluzione al problema è

$$\mathbf{f}_{m,n} = \sum_{k,l=0}^{N-1} \frac{\hat{g}_{k,l}}{\hat{K}_{k,l}} e^{\frac{2\pi i}{N}(km+ln)}.$$

Pertanto, una soluzione esiste e dipende con continuità dai dati. Nel caso in cui sia garantita l'unicità, il problema discreto è quindi ben posto nel senso di Hadamard.

Tuttavia, tale soluzione è in generale inaccettabile dal punto di vista fisico, in quanto è fortemente affetta da rumore.

Analizziamo questo fatto considerando una variazione $\delta\mathbf{g}$ dell'immagine discretizzata. Poichè la soluzione \mathbf{f} può essere espressa in termini della matrice inversa di \mathbf{A} come $\mathbf{f} = \mathbf{A}^{-1}\mathbf{g}$, la corrispondente variazione nella soluzione è data da

$$\delta\mathbf{f} = \mathbf{A}^{-1}\delta\mathbf{g}.$$

Sfruttando la submoltiplicatività della norma-2 si deduce che

$$\|\delta\mathbf{f}\| \leq \|\mathbf{A}^{-1}\| \|\delta\mathbf{g}\| = \frac{1}{\hat{K}_{\min}} \|\delta\mathbf{g}\|,$$

essendo $\hat{K}_{\min} = \min_{k,l} |\hat{K}_{k,l}|$.

D'altra parte, poichè $\mathbf{g} = \mathbf{A}\mathbf{f}$, vale anche che

$$\|\mathbf{g}\| \leq \hat{K}_{\max} \|\mathbf{f}\|,$$

con $\hat{K}_{\max} = \max_{k,l} |\hat{K}_{k,l}|$.

Rapportando membro a membro le ultime due relazioni si ottiene la disuguaglianza

$$\frac{\|\delta \mathbf{f}\|}{\|\mathbf{f}\|} \leq \frac{\hat{K}_{\max}}{\hat{K}_{\min}} \frac{\|\delta \mathbf{g}\|}{\|\mathbf{g}\|},$$

in cui il numero di condizionamento $\frac{\hat{K}_{\max}}{\hat{K}_{\min}}$ può essere elevato.

Il risultato ottenuto è tipico nei processi di discretizzazione per problemi di deconvoluzione qualunque e sottende un risultato dalla più generale validità, secondo il quale *nel discretizzare un problema mal posto il numero di condizionamento del corrispondente problema discreto è molto alto*.

Le proprietà numeriche che caratterizzano il problema del deblurring introducono la necessità di dover ricercare algoritmi e tecniche di regolarizzazione che pongano rimedio al mal condizionamento del problema e che consentano di determinare, per esso, delle soluzioni numericamente accettabili.

Questi temi saranno oggetto di approfondimento del prossimo capitolo.

Capitolo 4

Metodi di regolarizzazione

Nel capitolo precedente abbiamo visto che un modello numerico per il deblurring si presenta nella forma $A\mathbf{f} = \mathbf{g}$, dove A è una matrice mal condizionata, \mathbf{f} è l'immagine da ricostruire e \mathbf{g} è l'immagine sfocata. Per il mal condizionamento del problema, anche qualora la matrice A sia invertibile, la soluzione $\mathbf{f}_{\text{naive}} = A^{-1}\mathbf{g}$ è inaccettabile. Le tecniche numeriche di regolarizzazione consentono, attraverso l'aggiunta di opportune informazioni, di migliorare tale soluzione, qualificandosi, dunque, come metodi pratici atti alla risoluzione di problemi discreti mal condizionati.

4.1 I metodi di filtraggio spettrale

Esplicitando la presenza del rumore $\boldsymbol{\varepsilon}$, consideriamo il problema nella seguente formulazione:

$$A\mathbf{f} = \mathbf{g} + \boldsymbol{\varepsilon},$$

dove \mathbf{g} , in questo caso, denota l'immagine sfocata priva di rumore. Si ha dunque che

$$\mathbf{f}_{\text{naive}} = A^{-1}\mathbf{g} + A^{-1}\boldsymbol{\varepsilon}. \quad (4.1)$$

Una maggiore informazione sul termine $A^{-1}\boldsymbol{\varepsilon}$ può essere ottenuta utilizzando la decomposizione a valori singolari della matrice A .

Supponiamo che $A \in \mathbb{R}^{n \times n}$ e che tutti i valori singolari σ_i siano strettamente positivi. Sotto queste ipotesi A è invertibile e per l'ortogonalità delle matrici U e V si ha che

$$A^{-1} = (U\Sigma V^t)^{-1} = (V^t)^{-1}\Sigma^{-1}U^{-1} = V\Sigma^{-1}U^t,$$

dove Σ^{-1} è la matrice diagonale avente entrate $\frac{1}{\sigma_i}$ per $i = 1, \dots, n$.

Esplicitando, poi, le SVD della matrice A e della sua inversa A^{-1} si ottengono le

seguenti relazioni:

$$A = \begin{bmatrix} \mathbf{u}_1 & \mathbf{u}_2 & \dots & \mathbf{u}_n \end{bmatrix} \begin{bmatrix} \sigma_1 & & & \\ & \ddots & & \\ & & \ddots & \\ & & & \sigma_n \end{bmatrix} \begin{bmatrix} \mathbf{v}_1^t \\ \vdots \\ \vdots \\ \mathbf{v}_n^t \end{bmatrix} = \sum_{i=1}^n \sigma_i \mathbf{u}_i \mathbf{v}_i^t; \quad (4.2)$$

$$A^{-1} = \begin{bmatrix} \mathbf{v}_1 & \mathbf{v}_2 & \dots & \mathbf{v}_n \end{bmatrix} \begin{bmatrix} \frac{1}{\sigma_1} & & & \\ & \ddots & & \\ & & \ddots & \\ & & & \frac{1}{\sigma_n} \end{bmatrix} \begin{bmatrix} \mathbf{u}_1^t \\ \vdots \\ \vdots \\ \mathbf{u}_n^t \end{bmatrix} = \sum_{i=1}^n \frac{\mathbf{v}_i \mathbf{u}_i^t}{\sigma_i}. \quad (4.3)$$

Utilizzando la seconda delle precedenti, segue che la (4.1) può essere espressa interamente come combinazione lineare degli elementi della base $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ dei vettori singolari destri:

$$\mathbf{f}_{\text{naive}} = A^{-1} \mathbf{g} + A^{-1} \boldsymbol{\varepsilon} = V \Sigma^{-1} U^t \mathbf{g} + V \Sigma^{-1} U^t \boldsymbol{\varepsilon} = \sum_{i=1}^n \frac{\mathbf{u}_i^t \mathbf{g}}{\sigma_i} \mathbf{v}_i + \sum_{i=1}^n \frac{\mathbf{u}_i^t \boldsymbol{\varepsilon}}{\sigma_i} \mathbf{v}_i \quad (4.4)$$

Per comprendere quanto l'errore espresso dal termine

$$A^{-1} \boldsymbol{\varepsilon} = \sum_{i=1}^n \frac{\mathbf{u}_i^t \boldsymbol{\varepsilon}}{\sigma_i} \mathbf{v}_i \quad (4.5)$$

che appare nella (4.4) influenzi la soluzione, è necessario conoscere alcune proprietà che normalmente sussistono per problemi di deblurring:

- le componenti dell'errore, cioè le $\mathbf{u}_i^t \boldsymbol{\varepsilon}$, sono piccole in modulo e normalmente dello stesso ordine di grandezza per ogni $i = 1, \dots, n$;
- i valori singolari decrescono fino a valori prossimi allo zero; ciò è coerente col fatto che $\text{cond}(A) = \frac{\sigma_1}{\sigma_n}$ sia grande in modulo e, dunque, la soluzione sensibile a piccole perturbazioni ed errori di arrotondamento.
- I vettori singolari corrispondenti ai valori singolari più piccoli costituiscono le informazioni associate alle frequenze più alte, cioè a dire, al crescere del valore dell'indice i i vettori \mathbf{u}_i e \mathbf{v}_i tendono ad avere più cambi di segno.

Osserviamo inoltre, che per la (4.5), si ha che se le componenti $\frac{\mathbf{u}_i^t \boldsymbol{\varepsilon}}{\sigma_i}$ sono piccole in modulo per ogni $i = 1, \dots, n$, allora il termine $A^{-1} \boldsymbol{\varepsilon}$ influenza in modo trascurabile la soluzione espressa dalla (4.1); tuttavia, come accade nel caso del problema del deblurring, questa situazione non sempre si verifica nella pratica e i valori singolari tendono a decrescere fino a valori prossimi allo zero, mentre il rumore ha componenti più o meno uguali lungo tutti i vettori singolari sinistri; in questi casi l'informazione

relativa alle frequenze più alte risulta affetta da errori notevoli, rendendo la soluzione espressa dalla (4.4) distante da quella reale.

Per ovviare a questo problema, tra i metodi che possono essere utilizzati vi sono il *metodo della SVD troncata (TSVD)* e il *metodo di regolarizzazione di Tikhonov*, i quali appartengono alla più generale classe dei metodi di filtraggio spettrale.

Definizione.

Dato un sistema quadrato di ordine n ,

$$A\mathbf{x} = \mathbf{b},$$

con A mal condizionata e invertibile, un metodo di filtraggio spettrale consiste nella scelta per $i = 1, \dots, n$ di certi coefficienti ϕ_i , detti *fattori di filtro*, nella soluzione

$$\mathbf{x}_{\text{filt}} = \sum_{i=1}^n \phi_i \frac{\mathbf{u}_i^t \mathbf{b}}{\sigma_i} \mathbf{v}_i, \quad (4.6)$$

essendo $\{\mathbf{u}_1, \dots, \mathbf{u}_n\}$ e $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ rispettivamente i vettori singolari sinistri e destri di A .

4.1.1 La SVD troncata (TSVD)

Consideriamo un generico sistema

$$A\mathbf{x} = \mathbf{b},$$

e il corrispondente problema lineare ai minimi quadrati

$$\min_{\mathbf{x}} \|A\mathbf{x} - \mathbf{b}\|, \quad (4.7)$$

dove A è mal condizionata e avente δ -rango pari a k .

Il *teorema di Gauss-Markov* afferma che la soluzione ai minimi quadrati è il miglior stimatore lineare non distorto di \mathbf{x} , nel senso che è quello con varianza minima; nel nostro caso, in cui A è mal condizionata, tale minima varianza è comunque grande. Tuttavia, se si ammette che lo stimatore possa essere distorto, la varianza può essere drasticamente ridotta e un modo per ottenere tale risultato è troncare la soluzione ottenuta tramite la decomposizione a valori singolari, eliminando l'informazione associata alle più alte frequenze, cioè le componenti maggiormente affette da errore.

Si consideri, allora, la decomposizione a valori singolari di A ,

$$A = U\Sigma V^t = \sum_{i=1}^n \mathbf{u}_i \sigma_i \mathbf{v}_i^t.$$

Ponendo pari a zero tutti i valori singolari σ_i , $i > k$, per un certo k fissato, la soluzione corrispondente può essere espressa in termini dei primi k vettori singolari destri come

$$\mathbf{x}_{\text{filt}} = \sum_{i=1}^k \frac{\mathbf{u}_i^t \mathbf{b}}{\sigma_i} \mathbf{v}_i, \quad (4.8)$$

la quale viene detta *soluzione regolarizzata* tramite la *TSVD*. La soluzione individuata equivale a richiedere che i fattori di filtro nella (4.6) siano

$$\phi_i = \begin{cases} 1, & i = 1, \dots, k \\ 0, & i = k + 1, \dots, n. \end{cases}$$

Il parametro k viene detto *parametro di troncamento* e determina, dunque, il numero di componenti della *SVD* che vengono mantenute nella soluzione regolarizzata.

Osserviamo, inoltre, che la (4.7) risolve il problema ai minimi quadrati

$$\min_x \|A_k \mathbf{x} - \mathbf{b}\|, \quad A_k = \sum_{i=1}^k \mathbf{u}_i \sigma_i \mathbf{v}_i^t,$$

dove A_k costituisce la migliore approssimazione a rango k della matrice A .

Per la definizione di δ -rango segue, dunque, che

$$\|A - A_k\| = \|AV_2\| \leq \delta, \quad V_2 = \begin{bmatrix} \mathbf{v}_{k+1} & \dots & \mathbf{v}_n \end{bmatrix}$$

4.1.2 Regolarizzazione di Tikhonov

Supponiamo di voler determinare la soluzione per il problema

$$\min_x \|A\mathbf{x} - \mathbf{b}\|^2 + \tau^2 \|D\mathbf{x}\|^2, \quad (4.9)$$

dove $D = \text{diag}(d_1, \dots, d_n)$ è una matrice diagonale con entrate positive. Il problema (4.9) equivale al problema ai minimi quadrati

$$\min_{\mathbf{x}} \left\| \begin{bmatrix} \tau D \\ A \end{bmatrix} \mathbf{x} - \begin{bmatrix} 0 \\ \mathbf{b} \end{bmatrix} \right\|, \quad (4.10)$$

il quale, quando $\tau > 0$, è sempre a rango pieno e ha un'unica soluzione.

Se $D = I$ i valori singolari della matrice

$$\begin{bmatrix} \tau D \\ A \end{bmatrix}$$

che appare nella (4.10) sono uguali a

$$\tilde{\sigma}_i = \sqrt{\sigma_i^2 + \tau^2} \quad i = 1, \dots, n.$$

In questo caso la soluzione può essere espressa come combinazione lineare dei vettori singolari destri di A come

$$\mathbf{x}_\tau = \sum_{i=1}^n \frac{\sigma_i^2}{\sigma_i^2 + \tau^2} \frac{\mathbf{u}_i^\dagger \mathbf{b}}{\sigma_i} \mathbf{v}_i, \quad (4.11)$$

cioè a dire, scegliendo come fattori di filtro ϕ_i da sostituire nella (4.5) i valori

$$f_i = \frac{\sigma_i^2}{\sigma_i^2 + \tau^2} \quad i = 1, \dots, n.$$

Tali quantità sono dette *fattori di filtro di Tikhonov*.

Osservazione.

Sia $\sigma_1 < \tau < \sigma_n$. Allora vale che:

- se $\sigma_i \gg \tau$ si ha che $f_i = 1 + \mathcal{O}\left(\frac{\tau^2}{\sigma_i^2}\right) \approx 1$;
- se $\sigma_i \ll \tau$ si ha che $f_i = \frac{\sigma_i^2}{\tau^2} + \mathcal{O}\left(\frac{\sigma_i^4}{\tau^4}\right) \approx \frac{\sigma_i^2}{\tau^2}$;
- se $\sigma_i \simeq \tau$ allora $f_i \approx \frac{1}{2}$.

Dall'osservazione precedente deduciamo, dunque, che i fattori di filtro che maggiormente danno contributo alla soluzione (4.11) sono quelli relativi alle prime componenti, corrispondenti a valori singolari maggiori di τ , e che le componenti della SVD che corrispondono a valori singolari minori di τ contribuiscono in misura minima alla (4.11).

In virtù di queste considerazioni, possiamo dunque dedurre che la soluzione di Tikhonov e la soluzione mediante la TSVD sono pressochè equivalenti quando i parametri di troncamento k e τ sono scelti in modo tale che $\sigma_k \simeq \tau$.

Osservazione. La norma-2 della soluzione non è sempre la quantità migliore da includere nella formulazione di Tikhonov. Infatti, è stato testato che per certi problemi è meglio minimizzare la norma-2 di una quantità che approssima una derivata della soluzione.

Ad esempio, se si usa una regola di quadratura del punto medio per la discretizzazione, allora le approssimazioni della derivata prima e della derivata seconda sono date rispettivamente da $L_1 \mathbf{x}$ e $L_2 \mathbf{x}$, dove

$$L_1 = \frac{1}{n} \begin{bmatrix} -1 & 1 & & \\ & \ddots & \ddots & \\ & & -1 & 1 \end{bmatrix}, \quad L_2 = \frac{1}{n^2} \begin{bmatrix} 1 & -2 & 1 & & \\ & \ddots & \ddots & \ddots & \\ & & 1 & -2 & 1 \end{bmatrix},$$

essendo L_1 e L_2 rispettivamente di ordine $(n-1) \times n$ e $(n-2) \times n$.

Considerando quanto detto, si giunge alla seguente formulazione generale della regolarizzazione di Tikhonov:

$$\min_{\mathbf{x}} \|\mathbf{A}\mathbf{x} - \mathbf{b}\|^2 + \tau^2 \|L_j \mathbf{x}\|^2, \quad (4.12)$$

dove l'indice j indica l'ordine di derivazione.

Questo caso richiede l'uso della *decomposizione a valori singolari generalizzata* (GSVD).

4.1.3 TSVD e regolarizzazione di Tikhonov per problemi bidimensionali

Consideriamo un problema di deconvoluzione bidimensionale discretizzato in termini dei prodotti di Kronecker, cioè il problema nella forma (2.23):

$$(\bar{A} \otimes A)\text{vec}(F) = \text{vec}(G)$$

e calcoliamo la corrispondente soluzione tramite la TSVD. Sia, dunque, $\text{vec}(F_k)$ la soluzione troncata ottenuta dalla SVD, essendo k il parametro di troncamento corrispondente al numero di valori singolari di $\bar{A} \otimes A$ inclusi nella soluzione regolarizzata, e sia F_k la matrice che corrisponde alla soluzione regolarizzata del corrispondente problema bidimensionale. Allora, sfruttando la proprietà espressa dalla (2.20) e la relazione (4.3) si deduce che

$$\text{vec}(F_k) = \sum_{i,j} \frac{(\bar{\mathbf{u}}_j \otimes \mathbf{u}_i)^t}{\bar{\sigma}_j \sigma_i} \text{vec}(G)(\bar{\mathbf{v}}_j \otimes \mathbf{v}_i),$$

dove la sommatoria è intesa rispetto alle k più grandi quantità dei prodotti $\bar{\sigma}_j \sigma_i$.

A questo punto, usando la (2.19) si giunge all'espressione

$$F_k = \sum_{i,j} \frac{\mathbf{u}_i^t G \bar{\mathbf{u}}_j}{\sigma_i \bar{\sigma}_j} \mathbf{v}_i \bar{\mathbf{v}}_j^t, \quad (4.13)$$

la quale consente di determinare la soluzione della TSVD bidimensionale in funzione delle due SVD di A e \bar{A} .

In maniera analoga si può esprimere la soluzione di Tikhonov. In questo caso la regolarizzazione è data da

$$\min \left\{ \left\| (\bar{A} \otimes A)\text{vec}(F) - \text{vec}(G) \right\|^2 + \tau^2 \left\| \text{vec}(F) \right\|^2 \right\},$$

la quale, utilizzando la (2.23) diviene

$$\min \left\{ \left\| \text{vec}(A F \bar{A}^t) - \text{vec}(G) \right\|^2 + \tau^2 \left\| \text{vec}(F) \right\|^2 \right\},$$

la quale è equivalente a

$$\min \left\{ \left\| A F \bar{A}^t - G \right\|_{\text{F}}^2 + \tau^2 \left\| F \right\|_{\text{F}}^2 \right\},$$

essendo $\|\cdot\|_{\text{F}}$ la norma matriciale di Frobenius.

Questa formulazione risulta associata al problema ai minimi quadrati

$$\min \left\| \begin{bmatrix} \bar{A} \otimes A \\ \tau I \end{bmatrix} \text{vec}(F) - \begin{bmatrix} \text{vec}(G) \\ 0 \end{bmatrix} \right\|,$$

dove I è la matrice identità di ordine n^2 .

4.2 Criteri per la scelta del parametro di regolarizzazione

Alla luce dell'analisi effettuata per i metodi di regolarizzazione di Tikhonov e della SVD troncata, si comprende che, nella pratica, risulta necessaria la corretta scelta del parametro di regolarizzazione che determini il giusto equilibrio tra la presenza dell'errore di perturbazione dovuto al rumore e quella dell'errore prodotto dal filtraggio. Tale funzione viene eseguita da alcuni metodi specializzati, i quali, secondo quanto precisato in [13], possono essere distinti in due categorie principali:

1. metodi basati sulla conoscenza, o su una buona stima, di $\|\varepsilon\|$, con ε vettore associato al rumore;
2. metodi che non richiedono la conoscenza o la stima di $\|\varepsilon\|$.

Nei sottoparagrafi che seguono descriveremo brevemente un metodo rappresentante della prima delle due categorie, il *principio di discrepanza*, e due metodi appartenenti alla seconda, cioè a dire la *Generalized cross-validation (GCV)* e il *metodo della L-curva*.

Il principio di discrepanza

La scelta del parametro di regolarizzazione attraverso il principio di discrepanza si basa sulla possibilità di avere una buona stima di δ , con $\delta = \|\varepsilon\|$.

Se tale stima è nota allora il parametro di regolarizzazione può essere scelto in modo tale che la norma del residuo sia approssimativamente uguale a δ . Pertanto, la soluzione regolarizzata $\mathbf{x}_{\text{filt}}^*$ che, con il principio di discrepanza, viene considerata come essere la miglior stima è quella che realizza la relazione

$$\|\mathbf{b} - A\mathbf{x}_{\text{filt}}^*\| = \mu\delta, \quad (4.14)$$

per un certo $\mu \geq 1$ opportunamente definito (seguendo quanto precisato in [15] normalmente $\mu \in [2, 5]$).

Generalized cross-validation (GCV)

Nel caso del deblurring, le informazioni sul rumore sono spesso assenti e si deve, dunque, ricorrere a metodi che pongano rimedio a questa mancanza di informazione. Tra i principali candidati vi è il cosiddetto metodo basato sulla *Generalized cross-validation (GCV)*.

Tale metodo consente di determinare il parametro τ^* che minimizza una particolare funzione dove τ è il parametro di Tikhonov o, con abuso di notazione, $\tau = 1/k$,

con k parametro di troncamento nella TSVD. Si dimostra che una stima del miglior parametro per i metodi di filtraggio esaminati precedentemente minimizza il funzionale

$$G(\tau) = \frac{\|(I_n - AV\Phi\Sigma^{-1}U^t)\mathbf{b}\|_2^2}{(\text{tr}(I_n - AV\Phi\Sigma^{-1}U^t))^2}, \quad (4.15)$$

dove Φ è la matrice diagonale contenente i fattori di filtro, dipendenti da τ .

Sviluppando separatamente il numeratore e il denominatore della precedente espressione si deducono le seguenti:

$$\|(I_n - AV\Phi\Sigma^{-1}U^t)\mathbf{b}\|_2^2 = \|V - A\mathbf{x}_{\text{filt}}\|_2^2,$$

$$\text{tr}(I_n - AV\Phi\Sigma^{-1}U^t) = \text{tr}(I_n U \Sigma V^t V \Phi \Sigma^{-1} U^t) = \text{tr}(I_n - \Phi) = n - \sum_{i=1}^n \phi_i,$$

dalle quali segue che la (4.15) è equivalente a

$$G(\tau) = \frac{\|\mathbf{b} - A\mathbf{x}_{\text{filt}}\|_2^2}{(n - \sum_{i=1}^n \phi_i)^2}.$$

Metodo della L-curva

Il metodo della L-curva (cfr. [9], [14]) consiste nell'analisi della spezzata avente punti

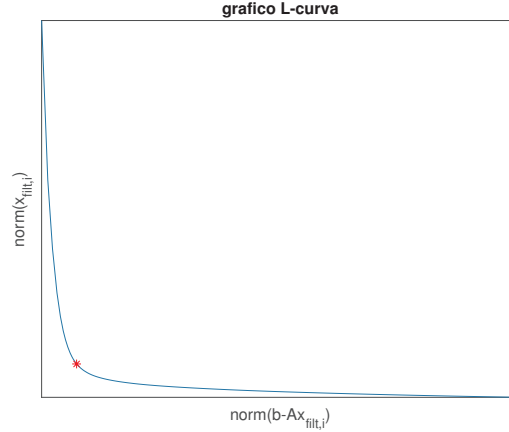
$$(x_i, y_i) = (\log_{10} \|A\mathbf{x}_{\text{filt},i} - \mathbf{b}\|_2, \log_{10} \|\mathbf{x}_{\text{filt},i}\|_2) \quad i = 1, \dots, p$$

dove p indica la cardinalità dell'insieme $\{\tau_1, \dots, \tau_p\}$ dei parametri di regolarizzazione e $\mathbf{x}_{\text{filt},i}$ indica la soluzione regolarizzata corrispondente a τ_i , calcolata, per esempio, attraverso la TSVD o il metodo di Tikhonov.

Nella maggior parte dei casi la spezzata che si ottiene presenta una caratteristica forma a "L" e il valore ottimale della regolarizzazione del parametro è quello corrispondente al punto di massima curvatura, cioè quello che maggiormente si avvicina all'angolo della "L". Questa scelta è giustificata dal fatto che:

- per valori del parametro maggiori la norma dei residui cresce rapidamente mentre la norma della soluzione regolarizzata decresce lentamente;
- per valori minori, si ha l'effetto contrario di quanto indicato al punto precedente, ovvero si verifica una crescita rapida nelle norme delle soluzioni e una corrispondente decrescita lenta nelle norme dei residui.

Ci si aspetta, dunque, che le coordinate del punto nei pressi dell'angolo della L-curva, corrispondente alla soluzione, siano quelle che realizzano, in modo ottimale, il $\min_i \|\mathbf{x}_{\text{filt},i}\|$ e il $\min_i \|\mathbf{b} - A\mathbf{x}_{\text{filt},i}\|$.



(a) Grafico della L-curva; in rosso: angolo della L-curva

4.3 Metodi iterativi di regolarizzazione

I metodi di regolarizzazione iterativi vengono normalmente applicati per problemi inversi di grandi dimensioni, in cui i metodi diretti hanno un costo computazionale troppo alto per poter essere applicati. Inoltre, dal momento che tali metodi chiamano in causa la matrice dei coefficienti soltanto per effettuare moltiplicazioni matrici-vettore, essi risultano particolarmente performanti nel caso di matrici sparse o strutturate, per le quali possono essere implementati gli *algoritmi fast*.

Riferendoci principalmente a [8], [3], [15], nei prossimi paragrafi vedremo come uno dei metodi iterativi più studiati, *il metodo del gradiente*, possa essere opportunamente adattato per determinare la soluzione per problemi discreti mal condizionati.

4.3.1 Metodo del gradiente

Sia A una matrice simmetrica e definita positiva e consideriamo la forma quadratica

$$\Psi(\mathbf{y}) = \frac{1}{2}\mathbf{y}^t A \mathbf{y} - \mathbf{y}^t \mathbf{b}, \quad (4.16)$$

la quale è minima nel punto in cui si annulla il suo gradiente

$$\nabla \Psi(\mathbf{y}) = \frac{1}{2} \nabla (\mathbf{y}^t A \mathbf{y}) - \nabla (\mathbf{y}^t \mathbf{b}) = 0. \quad (4.17)$$

Poichè per ogni k sussiste che

$$\begin{aligned} \frac{\partial}{\partial x_k} (\mathbf{x}^t \mathbf{b}) &= \frac{\partial}{\partial x_k} \left(\sum_{i=1}^n x_i b_i \right) = b_k, \\ \frac{\partial}{\partial x_k} (\mathbf{x}^t A \mathbf{x}) &= \frac{\partial}{\partial x_k} \left(\sum_{i,j=1}^n a_{ij} x_i x_j \right) = \sum_{i=1}^n a_{ik} x_i + \sum_{j=1}^n a_{kj} x_j = (A^t \mathbf{x})_k + (A \mathbf{x})_k, \end{aligned}$$

allora, sfruttando le ipotesi su A , la (4.17) diventa

$$\nabla\Psi(\mathbf{y}) = \frac{1}{2}(A + A^t)\mathbf{y} - \mathbf{b} = A\mathbf{y} - \mathbf{b} = 0.$$

Dunque, la ricerca del minimo della (4.16) corrisponde alla risoluzione del sistema lineare $A\mathbf{x} = \mathbf{b}$. Il metodo del gradiente consente di determinare la soluzione di tale sistema minimizzando $\Psi(\mathbf{y})$ attraverso un metodo iterativo del tipo *non stazionario*, cioè della forma

$$\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} + \alpha_k \mathbf{d}^{(k)},$$

a partire da un vettore iniziale $\mathbf{x}^{(0)}$, lungo delle direzioni prescelte determinate dai vettori $\mathbf{d}^{(k)}$ e con passi di lunghezza α_k .

Qualunque sia la direzione $\mathbf{d}^{(k)}$ è possibile determinare il minimo di $\Psi(\mathbf{x}^{(k+1)})$ rispetto alla variazione di $\alpha = \alpha_k$. Si osservi, innanzitutto, che

$$\begin{aligned} \Psi(\mathbf{x}^{(k+1)}) &= \frac{1}{2}(\mathbf{x}^{(k)} + \alpha \mathbf{d}^{(k)})^t A (\mathbf{x}^{(k)} + \alpha \mathbf{d}^{(k)}) - (\mathbf{x}^{(k)} + \alpha \mathbf{d}^{(k)})^t \mathbf{b} = \\ &= \frac{1}{2}(\mathbf{x}^{(k)})^t A \mathbf{x}^{(k)} + \alpha (\mathbf{d}^{(k)})^t A \mathbf{x}^{(k)} + \frac{1}{2} \alpha^2 (\mathbf{d}^{(k)})^t A \mathbf{d}^{(k)} - (\mathbf{x}^{(k)})^t \mathbf{b} - \alpha (\mathbf{d}^{(k)})^t \mathbf{b} = \\ &= \Psi(\mathbf{x}^{(k)}) - \alpha (\mathbf{d}^{(k)})^t \mathbf{r}^{(k)} + \frac{1}{2} \alpha^2 (\mathbf{d}^{(k)})^t A \mathbf{d}^{(k)}, \end{aligned}$$

dove è stato posto $\mathbf{r}^{(k)} = \mathbf{b} - A\mathbf{x}^{(k)}$.

Allora, imponendo nella relazione precedente che

$$\frac{d}{d\alpha} \Psi(\mathbf{x}^{(k+1)}) = 0$$

si ottiene che il passo ottimale è dato da

$$\alpha_k = \frac{(\mathbf{d}^{(k)})^t \mathbf{r}^{(k)}}{(\mathbf{d}^{(k)})^t A \mathbf{d}^{(k)}}. \quad (4.18)$$

Nel metodo del gradiente come direzione $\mathbf{d}^{(k)}$ si sceglie quelle di massima decrescita, ossia quella opposta alla direzione del gradiente, cioè:

$$\mathbf{d}^{(k)} = -\nabla\Psi(\mathbf{x}^{(k)}) = \mathbf{b} - A(\mathbf{x}^{(k)}) = \mathbf{r}^{(k)},$$

che, sostituito nella (4.18), consente di esprimere la lunghezza del passo ottimale come

$$\alpha_k = \frac{(\mathbf{r}^{(k)})^t \mathbf{r}^{(k)}}{(\mathbf{r}^{(k)})^t A \mathbf{r}^{(k)}}. \quad (4.19)$$

4.3.2 Metodo del gradiente coniugato

Il metodo del gradiente coniugato è una variante del metodo del gradiente ottenuta operando una scelta più accurata delle direzioni di discesa.

Per comprenderne meglio il funzionamento può essere utile illustrare i seguenti risultati.

Definizione. Un vettore $\mathbf{x}^{(k)}$ si dice *ottimale* rispetto a una direzione \mathbf{p} se

$$\Psi(\mathbf{x}^{(k)}) \leq \Psi(\mathbf{x}^{(k)} + \lambda\mathbf{p}), \quad \forall \lambda \in \mathbb{R}.$$

Teorema. Il vettore $\mathbf{x}^{(k)}$ è ottimale rispetto a \mathbf{p} se e solo se la direzione \mathbf{p} è ortogonale al residuo $\mathbf{r}^{(k)}$, ovvero se

$$\mathbf{p}^t \mathbf{r}^{(k)} = 0.$$

Dimostrazione.

Se $\mathbf{x}^{(k)}$ è ottimale rispetto a \mathbf{p} allora, per la definizione, la funzione $\Psi(\mathbf{x}^{(k)} + \lambda\mathbf{p})$ deve avere un minimo per $\lambda = 0$, cioè

$$\left. \frac{d}{d\lambda} \Psi(\mathbf{x}^{(k)} + \lambda\mathbf{p}) \right|_{\lambda=0} = 0. \quad (4.20)$$

Poichè

$$\begin{aligned} \Psi(\mathbf{x}^{(k)} + \lambda\mathbf{p}) &= (\mathbf{x}^{(k)} + \lambda\mathbf{p})^t A (\mathbf{x}^{(k)} + \lambda\mathbf{p}) - (\mathbf{x}^{(k)} + \lambda\mathbf{p})^t \mathbf{b} = \\ &= \Psi(\mathbf{x}^{(k)}) - \lambda \mathbf{p}^t (\mathbf{b} - A\mathbf{x}^{(k)}) + \frac{1}{2} \lambda^2 \mathbf{p}^t A \mathbf{p} = \\ &= \Psi(\mathbf{x}^{(k)}) - \lambda \mathbf{p}^t \mathbf{r}^{(k)} + \frac{1}{2} \lambda^2 \mathbf{p}^t A \mathbf{p}, \end{aligned}$$

la (4.20) implica che

$$\left. (\lambda \mathbf{p}^t A \mathbf{p} - \mathbf{p}^t \mathbf{r}^{(k)}) \right|_{\lambda=0} = 0,$$

ovvero che

$$\mathbf{p}^t \mathbf{r}^{(k)} = 0.$$

Viceversa, supponiamo che $\mathbf{p}^t \mathbf{r}^{(k)} = 0$. Allora si ha che

$$\Psi(\mathbf{x}^{(k)} + \lambda\mathbf{p}) = \Psi(\mathbf{x}^{(k)}) + \frac{1}{2} \lambda^2 \mathbf{p}^t A \mathbf{p}$$

e, dunque,

$$\frac{d}{d\lambda} \Psi(\mathbf{x}^{(k)} + \lambda\mathbf{p}) = \lambda \mathbf{p}^t A \mathbf{p}.$$

Usando la relazione precedente e il fatto che A è definita positiva segue che $\Psi(\mathbf{x}^{(k)} + \lambda\mathbf{p})$ ha un minimo per $\lambda = 0$. Pertanto:

$$\Psi(\mathbf{x}^{(k)}) \leq \Psi(\mathbf{x}^{(k)} + \lambda\mathbf{p}), \quad \forall \lambda \in \mathbb{R},$$

da cui la tesi. □

Osserviamo, a questo punto, che nel metodo del gradiente, per la particolare scelta del parametro α_k nella (4.19), il residuo $\mathbf{r}^{(k+1)}$ è ortogonale a $\mathbf{r}^{(k)}$. Infatti

$$\begin{aligned} (\mathbf{r}^{(k+1)})^t \mathbf{r}^{(k)} &= (\mathbf{b} - A\mathbf{x}^{(k+1)})^t \mathbf{r}^{(k)} = (\mathbf{b} - A(\mathbf{x}^{(k)} + \alpha_k \mathbf{r}^{(k)}))^t \mathbf{r}^{(k)} = \\ &= (\mathbf{r}^{(k)})^t \mathbf{r}^{(k)} - \alpha_k (\mathbf{r}^{(k)})^t A \mathbf{r}^{(k)} = (\mathbf{r}^{(k)})^t \mathbf{r}^{(k)} - \frac{(\mathbf{r}^{(k)})^t \mathbf{r}^{(k)}}{(\mathbf{r}^{(k)})^t A \mathbf{r}^{(k)}} (\mathbf{r}^{(k)})^t A \mathbf{r}^{(k)} = 0. \end{aligned}$$

Tuttavia, questo risultato non è più vero per il vettore $\mathbf{x}^{(k+2)}$ e per i successivi. Per fare in modo che venga mantenuta l'ottimalità del vettore k -esimo, rispetto a una certa direzione, anche alle iterazioni successive, supponiamo che $\mathbf{x}^{(k)}$ sia ottimale rispetto a \mathbf{p} e consideriamo iterazioni del tipo

$$\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} + \mathbf{q}.$$

Richiediamo, dunque, che $\mathbf{x}^{(k+1)}$ sia ottimale rispetto a \mathbf{p} , cioè

$$0 = \mathbf{p}^t \mathbf{r}^{(k+1)} = \mathbf{p}^t (\mathbf{b} - A\mathbf{x}^{(k)} - A\mathbf{q}) = \mathbf{p}^t \mathbf{r}^{(k)} - \mathbf{p}^t A\mathbf{q} = -\mathbf{p}^t A\mathbf{q}.$$

Sotto questa condizione, le due direzioni \mathbf{p} e \mathbf{q} si dicono *A-coniugate*.

Allora, partendo dalla direzione $\mathbf{p}^{(0)} = \mathbf{r}^{(0)}$ e considerando iterazioni della forma $\mathbf{p}^{(k+1)} = \mathbf{r}^{(k+1)} - \beta_k \mathbf{p}^{(k)}$, se si impone che $\mathbf{p}^{(k)}$ e $\mathbf{p}^{(k+1)}$ siano *A-coniugate* è possibile ottenere l'espressione di β_k :

$$0 = (\mathbf{p}^{(k)})^t A \mathbf{p}^{(k+1)} = (\mathbf{p}^{(k)})^t A (\mathbf{r}^{(k+1)} - \beta_k \mathbf{p}^{(k)}) = (\mathbf{p}^{(k)})^t A \mathbf{r}^{(k+1)} - \beta_k (\mathbf{p}^{(k)})^t A \mathbf{p}^{(k)},$$

da cui

$$\beta_k = \frac{(\mathbf{p}^{(k)})^t A \mathbf{r}^{(k+1)}}{(\mathbf{p}^{(k)})^t A \mathbf{p}^{(k)}}. \quad (4.21)$$

Effettuando questa scelta per il parametro β_k è possibile dimostrare che

$$(\mathbf{p}^{(i)})^t A \mathbf{p}^{(k+1)} = 0 \quad \forall i = 0, \dots, k,$$

cioè la direzione $\mathbf{p}^{(k+1)}$ è *A-coniugata* a tutte le direzioni generate ai passi precedenti.

4.3.3 Metodo del gradiente coniugato per problemi ai minimi quadrati (CGLS)

Se la matrice dei coefficienti A nel sistema $A\mathbf{x} = \mathbf{b}$ non è simmetrica e definita positiva, allora né il metodo del gradiente né il metodo del gradiente coniugato sono applicabili.

Tuttavia entrambi i metodi possono essere implementati per la risoluzione del sistema normale

$$A^t A \mathbf{x} = A^t \mathbf{b}$$

associato al problema ai minimi quadrati

$$\min \|A\mathbf{x} - \mathbf{b}\|_2.$$

Si parla, in questo caso, di *metodo del gradiente coniugato per minimi quadrati (CGLS)*.

Denotando con $\mathbf{x}^{(0)}$ la stima iniziale della soluzione e definendo i vettori ausiliari $\mathbf{r}^{(0)} = \mathbf{b} - A\mathbf{x}^{(0)}$ e $\mathbf{d}^{(0)} = A^t\mathbf{r}^{(0)}$, proponiamo la descrizione dell'algoritmo presente in [15].

Per $k = 1, 2, \dots$

$$\begin{aligned}\alpha_k &= \frac{\|A^t\mathbf{r}^{(k-1)}\|_2^2}{\|A\mathbf{d}^{(k-1)}\|_2^2}, \\ \mathbf{x}^{(k)} &= \mathbf{x}^{(k-1)} + \alpha_k\mathbf{d}^{(k-1)}, \\ \mathbf{r}^{(k)} &= \mathbf{r}^{(k-1)} - \alpha_k A\mathbf{d}^{(k-1)}, \\ \beta_k &= \frac{\|A^t\mathbf{r}^{(k)}\|_2^2}{\|A^t\mathbf{r}^{(k-1)}\|_2^2}, \\ \mathbf{d}^{(k)} &= A^t\mathbf{r}^{(k)} + \beta_k\mathbf{d}^{(k-1)},\end{aligned}$$

dove a ciascuna iterazione vengono calcolati il vettore dei residui per il problema ai minimi quadrati, $\mathbf{r}^{(k)} = \mathbf{b} - A\mathbf{x}^{(k)}$, e il vettore dei residui per le equazioni normali $A^t\mathbf{r}^{(k)} = A^t\mathbf{b} - A^tA\mathbf{x}^{(k)}$.

4.3.4 CGLS per problemi in dimensione 2

L'algoritmo appena proposto può essere facilmente esteso al caso di problemi bidimensionali, cioè aventi per incognite matrici, sfruttando il fatto che tali problemi possono essere espressi mediante un sistema della forma

$$(\overline{A} \otimes A)\text{vec}(F) = \text{vec}(G),$$

che, sfruttando la (2.22), risulta a sua volta equivalente a

$$AF\overline{A}^t = G.$$

Allora, riscrivendo le condizioni iniziali come

$$F^{(0)} = \text{stima iniziale della soluzione}$$

$$R^{(0)} = G - AF^{(0)}\overline{A}^t$$

$$D^{(0)} = A^t R^{(0)} \bar{A}$$

e sfruttando il fatto che $\|\text{vec}(X)\|_2 = \|X\|_F$, l'algoritmo precedente può essere opportunamente riscritto per essere adattato al caso bidimensionale:

$$\alpha_k = \frac{\|A^t R^{(k-1)} \bar{A}\|_F^2}{\|AD^{(k-1)} \bar{A}^t\|_F^2},$$

$$F^{(k)} = F^{(k-1)} + \alpha_k D^{(k-1)},$$

$$R^{(k)} = R^{(k-1)} - \alpha_k AD^{(k-1)} \bar{A}^t,$$

$$\beta_k = \frac{\|A^t R^{(k)} \bar{A}\|_F^2}{\|A^t R^{(k-1)} \bar{A}\|_F^2},$$

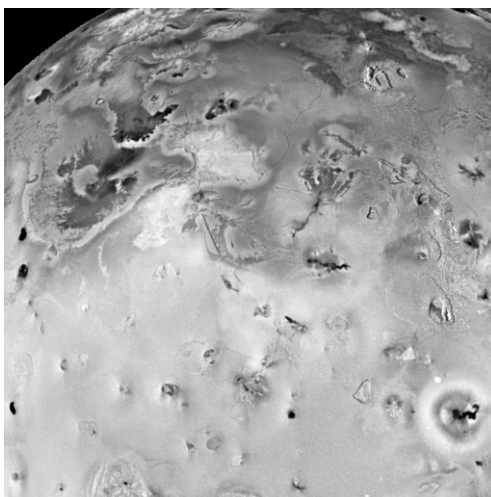
$$D^{(k)} = A^t R^{(k)} \bar{A} + \beta_k D^{(k-1)}.$$

Capitolo 5

Test numerici e sperimentazioni

Per lo sviluppo della parte sperimentale dell'elaborato, che verrà descritta in questo capitolo, sono stati consultati i pacchetti Matlab “HNO functions” e *Regularization Tools*, realizzati dal prof. Christian Hansen, i quali contengono un set di routines che consentono di determinare soluzioni regolarizzate per problemi discreti mal posti (cfr. [14],[15],[13]). Tra le funzioni presenti, alcune di quelle che sono state effettivamente utilizzate sono state opportunamente adattate agli scopi del presente lavoro, e, dunque, eventualmente modificate.

Consideriamo la seguente immagine test, che indicheremo con X , rappresentante una parte della superficie di un corpo celeste e avente dimensioni 512×512 (il corrispondente vettore $\text{vec}(X)$ ha quindi $512^2 = 262144$ componenti):



(a) rappresentazione immagine nitida (X)

La corrispondente immagine sfocata, denotata con B , su cui è stata effettuata la

sperimentazione, viene ottenuta sfocando X con una PSF a variabili separabili di tipo gaussiano e ammettendo che sia affetta da rumore bianco di tipo gaussiano, ε , soddisfacente la condizione $\frac{\|\varepsilon\|}{\|AX\|} = 0.05$.

Lo scopo della sperimentazione è stato, dunque, quello di ripristinare l'immagine danneggiata B utilizzando i metodi di regolarizzazione diretti della TSVD e di Tikhonov e il metodo iterativo CGLS, per ottenere un'immagine che fosse quanto più possibile simile all'originale.

5.1 I dati del problema: costruzione dell'immagine sfocata

Per poter giungere alla definizione dell'immagine sfocata B , occorre inizialmente costruire la matrice associata alla PSF per sfocatura di tipo gaussiano. È stata quindi definita una funzione opportuna, nominata `myPsfGauss`, la quale:

- richiede in input il numero di righe e colonne della matrice e i valori s_1 e s_2 delle deviazioni standard σ e $\bar{\sigma}$ nella (3.9);
- restituisce in output la matrice PSF e il punto `center` corrispondente al centro della Point Spread Function.

```
function[PSF,center]=myPsfGauss(m,n,s1,s2)
x=-fix(n/2):ceil(n/2)-1;
y=fix(m/2):ceil(m/2)-1; [X,Y]=meshgrid(x,y);
PSF=exp(-(X.^ 2)/(2*s1 ^ 2)-(Y.^ 2)/(s1*s2 ^ 2))/(2*pi*s1*s2);
[mm,nn]=find(PSF==max(PSF(:)));
center=[mm(1),nn(1)];
```

Il problema successivo che è stato necessario affrontare per costruire un buon modello di blurring, è quello legato alla scelta della boundary condition per la particolare immagine.

Limitandoci a considerare le tre tipologie illustrate nel Capitolo 3 l'imposizione della corretta boundary condition per poter ottenere un modello di sfocatura adatto può essere effettuata immergendo la matrice X all'interno di una matrice di dimensioni maggiori. In particolare:

1. se si assume la zero boundary condition l'immagine X viene immersa nella

matrice

$$X_{\text{ext}} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & X & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

2. se la condizione è di tipo periodico allora

$$X_{\text{ext}} = \begin{bmatrix} X & X & X \\ X & X & X \\ X & X & X \end{bmatrix}$$

3. per la condizione di tipo riflessivo la matrice estesa è, invece, della forma

$$X_{\text{ext}} = \begin{bmatrix} X_x & X_{\text{ud}} & X_x \\ X_{\text{lr}} & X & X_{\text{lr}} \\ X_x & X_{\text{ud}} & X_x \end{bmatrix},$$

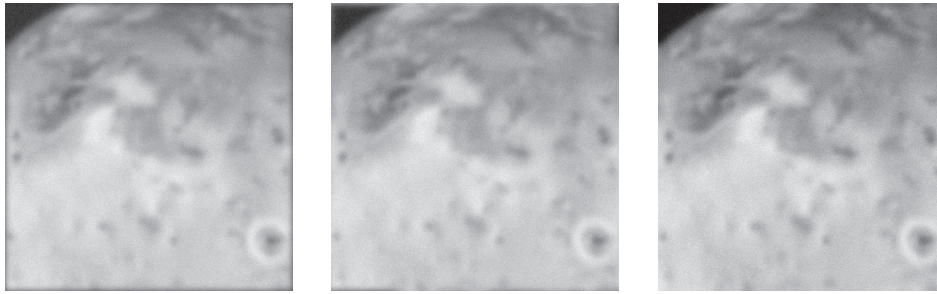
dove $X_{\text{ud}} = \text{flipud}(X)$, $X_{\text{lr}} = \text{fliplr}(X)$ e $X_x = \text{fliplr}(X_{\text{ud}})$ essendo flipud (*flip up to down*) e fliplr (*flip left to right*) le funzioni predefinite in Matlab che capovolgono l'ordine delle righe e delle colonne di una matrice rispettivamente dall'alto verso il basso e da sinistra verso destra.

Una volta definita l'estensione della matrice X , la matrice B che rappresenta l'immagine sfocata viene estratta dalla matrice B_{ext} ottenuta dalla convoluzione tra le matrici X_{ext} e PSF. Si aggiunge, poi, alla matrice sfocata, il contributo del rumore gaussiano.

Considerando quanto descritto, è stato implementato il seguente set di istruzioni:

```
B_ext=conv2(X_ext,PSF, 'same');
B= B_ext(513:1024,513:1024);
E=randn(size(B));
E=E/norm(E, 'fro');
B=B+0.05*norm(B, 'fro')*E;
```

Imponendo di volta in volta le tre diverse boundary conditions, è stato possibile confrontare i seguenti modelli per l'immagine sfocata:

(a) *Sfocatura con zero boundary condition*(b) *Sfocatura con periodic boundary condition*(c) *Sfocatura con reflexive boundary condition*

Effettuando uno zoom sui bordi si evince, infine, che la condizione che genera gli effetti meno indesiderati sui bordi dell'immagine è quella riflessiva; pertanto, è stata assunta quest'ultima condizione come quella da imporre sui dati a disposizione.

(a) *Zoom in zero boundary condition*(b) *Zoom in periodic boundary condition*(c) *Zoom in reflexive boundary condition*

5.2 Determinazione della soluzione numerica mediante metodi di regolarizzazione diretti

A questo punto, ci si è interessati alla formulazione del problema nella forma matriciale

$$(A_r \otimes A_c) \text{vec}(X) = \text{vec}(B)$$

dove A_r e A_c sono le matrici componenti del prodotto tensoriale di Kronecker che esprime la matrice dei coefficienti del sistema. Sfruttando nuovamente la (2.19), si osservi che la relazione precedente è equivalente alla seguente:

$$\text{vec}(A_c X A_r^t) = \text{vec}(B),$$

dalla quale segue che

$$A_c X A_r^t = B. \quad (5.1)$$

Dal punto di vista teorico la (5.1) consente, dunque, di esprimere la soluzione nella forma

$$X = A_c^{-1} B A_r^{-t}.$$

Allora, usando le SVD $A_c = U_c \Sigma_c V_c^t$, e $A_r = U_r \Sigma_r V_r^t$ si ottiene in conclusione che la soluzione è

$$X = (V_c \Sigma_c^{-1} U_c^t) B (U_r \Sigma_r^{-1} V_r^t), \quad (5.2)$$

la quale, poichè numericamente inaccettabile, dovrà essere opportunamente regolarizzata.

In sintesi, il procedimento adottato per la risoluzione del problema di image restoration per l'immagine test scelta è stato suddiviso nelle tre fasi seguenti:

1. determinazione delle matrici A_r e A_c che costituiscono la fattorizzazione di Kronecker;
2. individuazione della soluzione X , non accettabile numericamente;
3. regolarizzazione della soluzione X e conseguente determinazione della soluzione X_{filt} attraverso i metodi di regolarizzazione analizzati nel Capitolo 4.

5.2.1 Decomposizione di Kronecker

Le matrici A_r e A_c sono state determinate attraverso la funzione `mykronDecomp`, la quale assume in input la matrice associata alla PSF e il suo centro.

```
function [Ar,Ac]=mykronDecomp(PSF,center)
```

```
[U,S,V]=svd(PSF);
```

Poichè la *PSF* ha entrate non negative, i vettori della sua decomposizione a rango uno (cfr. (4.2)) devono avere componenti non negative. Per comprendere come rendere soddisfatta questa condizione è opportuno citare la seguente proprietà (cfr. [15]):

Proprietà. Per le matrici che provengono dalla discretizzazione di equazioni integrali di Fredholm del primo tipo, i vettori singolari \mathbf{u}_i e \mathbf{v}_i hanno un numero crescente di cambi di segno nelle proprie entrate al crescere di i , cioè in corrispondenza della decrescita dei valori singolari. Il numero di cambi di segno è precisamente $i - 1$.

Da questa proprietà segue che le prime colonne delle matrici U e V hanno entrate o tutte positive o tutte negative.

Pertanto, affinché la PSF abbia entrate non negative, è sufficiente richiedere che i vettori singolari corrispondenti al valore singolare più grande abbiano entrate non negative. Occorre, dunque, accertarsi che questa condizione sia rispettata, cambiando il segno qualora necessario.

```
minU=abs(min(U(:,1)));
maxU=max(abs(U(:,1)));
if minU==maxU
U=-U;
V=-V;
```

Poichè è stata imposta la boundary condition riflessiva, sappiamo, per quanto esposto nel paragrafo 3.3, che la struttura delle matrici A_r e A_c è data dalla somma di una matrice di Toeplitz e una matrice di Hankel e risultano definite, rispettivamente, dai vettori c e r :

```
c=sqrt(S(1,1))*U(:,1);
r=sqrt(S(1,1))*V(:,1);
A_r=buildToeplitz(r,center(2))+buildHankel(r,center(2));
A_c=buildToeplitz(c,center(1))+buildHankel(c,center(1));
```

dove le funzioni ausiliarie `buildToeplitz(c,k)` e `buildHankel(c,k)` generano rispettivamente la matrice di Toeplitz e la matrice di Hankel da una colonna centrale c , avente indice k :

```
function T=buildToeplitz(c,k)
n=length(c);
col=zeros(n,1);
row=col';
col(1:n-k+1,1)=c(k:n);
row(1,1:k)=c(k:-1:1)';
T=toeplitz(col,row);
```

```
function H=buildHankel(c,k)
n=length(c);
col=zeros(n,1);
```

```

col(1:n-k)=c(k+1:n);
row=zeros(n,1);
row(n-k+2:n)=c(1:k-1)
H=hankel(col,row);

```

5.2.2 Regolarizzazione tramite la TSVD

Osserviamo, che la relazione (5.2) consente di esprimere la soluzione X in termini delle fattorizzazioni a valori singolari delle matrici A_r e A_c :

$$X = V_c \Sigma_c^{-1} U_c^t B U_r \Sigma_r^{-1} V_r^t = V_c \Sigma_c^{-1} \hat{B} \Sigma_r^{-1} V_r^t,$$

dove si è posto $\hat{B} = U_c^t B U_r$.

Allora, considerando la soluzione regolarizzata mediante la TSVD e scrivendo in forma matriciale la (4.13), si ottiene che la soluzione filtrata può essere scritta come

$$X_{\text{filt}} = V_c (\hat{B} .* S_{\text{filt}}) V_r^t, \quad (5.3)$$

dove si ricorda che “.*” indica il prodotto di Hadamard-Schur e S_{filt} è la matrice di ordine 512×512 avente entrate

$$[S_{\text{filt}}]_{ij} = \begin{cases} 0 & \text{se } [\Sigma_c]_i [\Sigma_r]_j < \text{tol} \\ 1/[\Sigma_c]_i [\Sigma_r]_j & \text{se } [\Sigma_c]_i [\Sigma_r]_j \geq \text{tol} \end{cases},$$

essendo tol un parametro di tolleranza arbitrariamente scelto dall'utente.

La soluzione regolarizzata può allora essere ottenuta sperimentalmente implementando il codice

```

function [Xfilt,k]=mytsvd(B,PSF,center,tol)

[Ar,Ac]=mykronDecomp(PSF,center);
[Ur,Sr,Vr]=svd(Ar);
[Uc,Sc,Vc]=svd(Ac);
S=diag(Sc)*diag(Sr)';
Phi=(abs(S)>=tol);
idx=(S~=0);
Sfilt=zeros(size(Phi));
Sfilt(idx)=Phi(idx)./S(idx);
Bhat=Uc'*B*Ur;

```

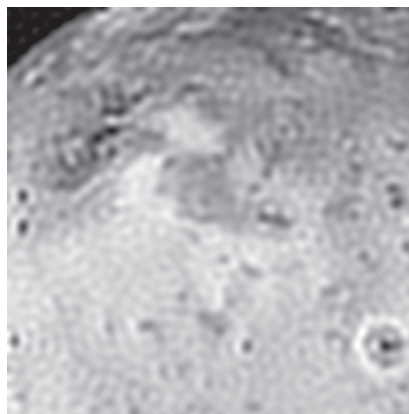
```
Xfilt=Vc*(Bhat.*Sfilt)*Vr';  
Phi=Phi(:);  
k=0;  
for j=1:512^2  
if Phi(j,1)==1  
k=k+1;  
end  
end
```

che, oltre alla soluzione filtrata, restituisce in output anche il parametro di troncamento k , cioè il numero di valori singolari maggiori o uguali al valore soglia tol indicato dall'utente.

A questo punto, per poter risalire alla soluzione regolarizzata occorre attribuire in input il valore del parametro tol , che può essere scelto applicando uno dei metodi descritti nel paragrafo 4.2.

Scelta del parametro usando la GCV

La funzione `gcvtsvd` presente nel pacchetto “HNO” consente di ottenere direttamente il parametro tol che minimizza la (4.15), che, nel caso in esame, vale 0.0780. In corrispondenza di questo valore del parametro, la SVD viene troncata dopo 3020 termini e si ottiene la seguente soluzione:



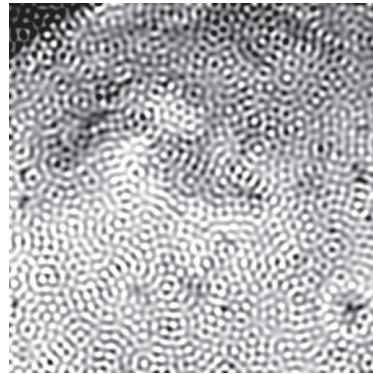
(a) *Soluzione regolarizzata mediante la TSVD con scelta del parametro effettuata usando la GCV*

Scelta del parametro usando il metodo della L-curva

Per stimare il parametro di regolarizzazione ottimale utilizzando il *metodo della L-curva*, è stato eseguito il grafico in scala logaritmica dei punti aventi coordinate (x, y) pari alle entrate dei vettori **rho** ed **eta** corrispondenti, rispettivamente, alle norme dei residui $\|\mathbf{b} - \mathbf{A}\mathbf{x}_{\text{filt},i}\|$ e delle soluzioni regolarizzate $\|\mathbf{x}_{\text{filt},i}\|$ con $i = 1, \dots, n$ per un certo n opportuno. Le soluzioni $\{\mathbf{x}_{\text{filt},1}, \mathbf{x}_{\text{filt},2}, \dots, \mathbf{x}_{\text{filt},n}\}$ sono state calcolate chiamando la funzione `mytsvd` assegnando in input n valori distinti del parametro `tol`.

Per comprendere quali valori attribuire a `tol`, è stato inizialmente osservato che in corrispondenza di valori dell'ordine del millesimo si ottenevano immagini degradate per l'eccessiva presenza di rumore.

Per `tol=0.008`, cioè arrestando la SVD dopo 5686 termini, si è ottenuto, ad esempio, il seguente risultato:



(a) Immagine degradata ottenuta per `tol=0.008`.

Allora, scegliendo questo valore come quantità da incrementare per ottenere la successione delle corrispondenti soluzioni regolarizzate, è stato possibile definire i vettori **rho** ed **eta** definendo la funzione seguente:

```
function [eta,rho]= norm_eta_rho(B,PSF,center,n)
sol_filt=zeros(n*512,512);
eta=zeros(n,1);
rho=zeros(n,1);
for i=1:n
tol=0.0080+i*0.003;
solreg=mytsvd(B,PSF,center,tol);
sol_filt(1+512*(i-1):i*512,1:512)=solreg;
rho(i)=norm(B(:)-prodottoAvett(PSF,center,solreg(:)));
```



```
eta(i)=norm(solreg,'fro');
end
```

dove la funzione `prodottoAvett` esegue il prodotto matrice-vettore sfruttando la struttura della matrice dei coefficienti, la quale, essendo del tipo $BTTB + BTHB + BHTB + BHHB$ può essere fattorizzata come

$$A = C^t \Lambda C,$$

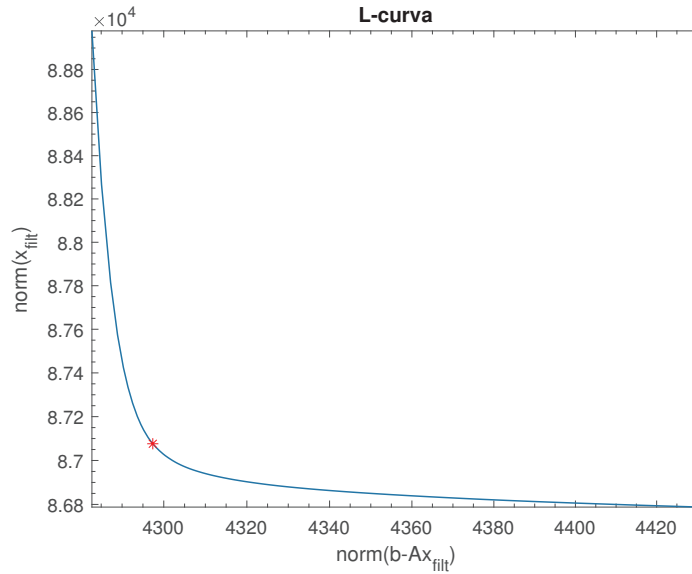
dove C è la matrice associata alla DCT2.

```
function pr= prodottoAx(PSF,center,x)

%elementi dello spettro della matrice A=C'AC, dove C è la matrice associata
alla dct2
e1=zeros(size(PSF));
e1(1,1)=1;
S=dct2(dctshift(PSF,center))./dct2(e1);
%Calcolo del prodotto sfruttando la struttura
X=reshape(x,size(PSF));
pr=idct2(S.*dct2(X));
pr=pr(:);
```

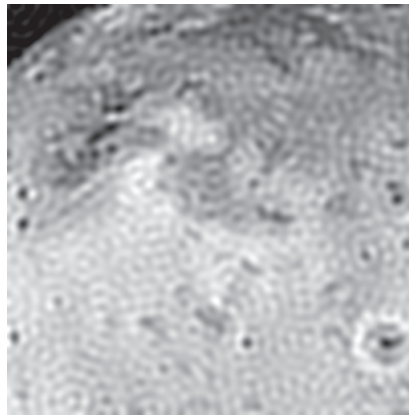
Osserviamo che nel codice appena proposto viene chiamata la funzione `dctshift` presente nel pacchetto “HNO”, la quale restituisce in output l’array contenente la prima colonna della matrice di blurring.

Per un n opportuno, il grafico della L-curva corrispondente evidenzia che il punto più vicino all’angolo della curva è ottenuto alla tredicesima iterazione del ciclo `for` implementato nella funzione `norm_eta_rho`, in corrispondenza della quale si ha `tol=0.047`.



(a) Grafico della *L*-curva per $n = 100$; in rosso il punto $(\rho(13), \eta(13))$

Si ha, dunque, la seguente soluzione:



(a) Soluzione regolarizzata mediante la TSVD con scelta del parametro effettuata usando il metodo della *L*-curva

Scelta del parametro usando il principio di discrepanza

Per scegliere il parametro di regolarizzazione utilizzando il principio di discrepanza è stata definita una funzione che determinasse il

$$\min_{X_{\text{filt},i}} \left\{ \left| \|B - AX_{\text{filt},i}\|_F - \mu\sigma \|B\|_F \right| \right\}, \quad (5.4)$$

dove:

1. B è la matrice contenente l'informazione relativa alla sfocatura;

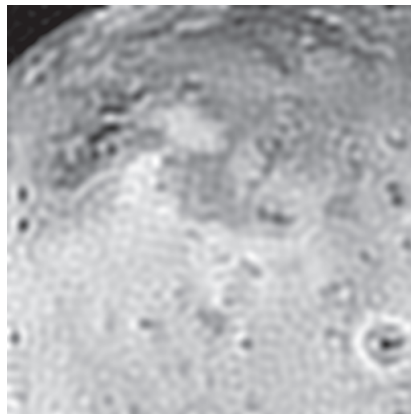
2. $\sigma = 0.05$ indica la deviazione standard del rumore, cioè $\frac{\|E\|_F}{\|B\|_F} = 0.05$;
3. $\mu \in [2, 5]$ è il parametro nella (4.14), a cui viene attribuito il valore $\mu = 3.5$.

Con questi dati a disposizione, in maniera analoga a quanto fatto nel caso del metodo della L-curva, si implementa una funzione che calcoli i residui corrispondenti alle soluzioni $\{X_{\text{filt},1}, X_{\text{filt},2}, \dots, X_{\text{filt},n}\}$:

```
function rho_discr=norm_rho_discr(B,PSF,center,n)

sol=zeros(n*512,512);
rho_discr=zeros(n,1);
for i=1:n
tol=0.02+i*0.002;
solreg=mytsvd_sep(B,PSF,center,tol);
soluzioni(1+512*(i-1):i*512,1:512)=solreg;
rho_discr(i)=norm(B(:)-prodottoAvett(PSF,center,solreg(:)));
end
```

Infine, ricercando con il comando `find` le componenti di `rho_discr` minori di $\mu\sigma\|B\|_F$, è stato possibile stabilire che il valore del parametro `tol` suggerito dall'applicazione del principio di discrepanza vale 0.072 (dopo 26 iterazioni del ciclo `for` nella function `rho_discr`), in corrispondenza del quale la `mytsvd` restituisce il seguente risultato:



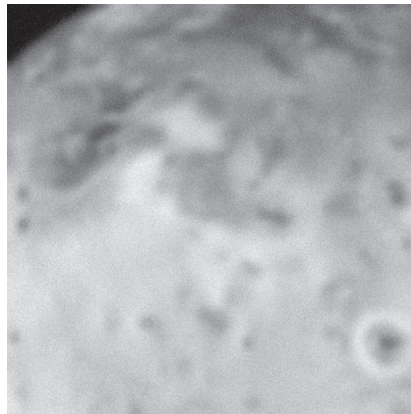
(a) *Soluzione regolarizzata mediante la TSVD con scelta del parametro effettuata usando il principio di discrepanza.*

Confronto tra le soluzioni

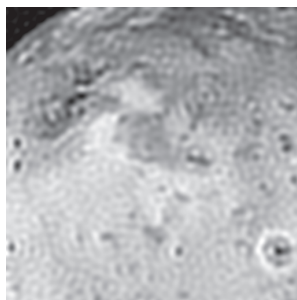
Riassumiamo nella tabella sottostante i risultati ottenuti dall'applicazione della TSVD, indicando il metodo utilizzato per la scelta di parametro `tol` e il numero di componenti k considerate nella (4.8).

Metodo	<code>tol</code>	k
GCV	0.078	3020
L-curva	0.047	3610
Pr. di discrepanza	0.072	3117

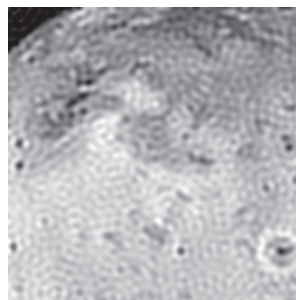
Per quanto diverse, tutte e tre le soluzioni ottenute costituiscono un chiaro miglioramento dell'immagine di partenza, acquisita come dato.



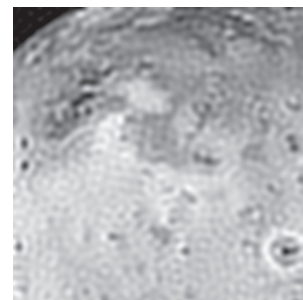
(a) Immagine sfocata iniziale



(a) Sol. GCV



(b) Sol. metodo della L-curva



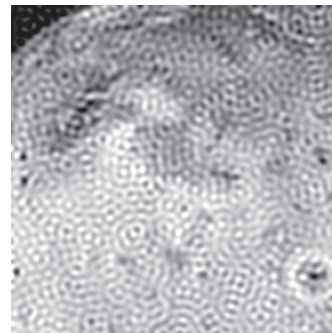
(c) Sol. principio di discrepanza

Osserviamo, inoltre, che attribuendo al parametro `tol` valori superiori dei tre valori dedotti si ottengono immagini meno affette da rumore, ma in cui viene persa più informazione; viceversa, assegnando valori inferiori, si ottengono immagini che,

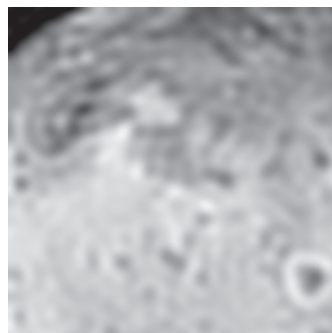
seppur contenenti più informazione, risultano degradate per l'eccessiva presenza di rumore nella soluzione.



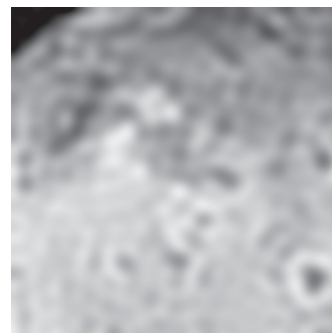
(a) *output per tol=0.01;*
k=5423



(b) *output per tol=0.02;*
k=4604



(c) *output per tol=0.4; k=1099*



(d) *output per tol=0.5; k=833*

Quanto osservato suggerisce che la scelta del parametro di troncamento tramite i criteri di scelta utilizzati ha consentito di determinare una soluzione che bilanci una parziale riduzione del rumore con una controllata perdita di informazione.

5.2.3 Regolarizzazione tramite filtraggio di Tikhonov

In maniera simile a quanto visto per la determinazione della soluzione regolarizzata con la TSVD è possibile determinare la stessa mediante il metodo di Tikhonov:

```
function[Xfilttick,tau]=tikhonov(B,PSF,center,tau)
```

```
[Ar,Ac]=mykronDecomp(PSF,center);
```

```
[Uc,Sc,Vc]=svd(Ac);
```

```
[Ur,Sr,Vr]=svd(Ar);
```

```
bhat=Uc'*B*Ur;
```

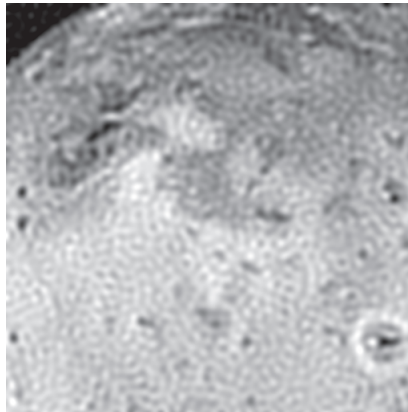
```
bhat=bhat(:);
```

```

s=kron(diag(Sr),diag(Sc));
D=abs(s).^ 2+abs(tau)^ 2;
bhat=s.*bhat;
xhat=bhat./D;
xhat=reshape(xhat,size(B));
Xfilttick=Vc*xhat*Vr';

```

Utilizzando come criterio per la scelta del parametro di troncamento la Generalized cross-validation si perviene al risultato seguente:



(a) Soluzione regolarizzata tramite il metodo di Tikhonov con uso della GCV per la scelta del parametro

5.3 Determinazione della soluzione numerica mediante metodi di regolarizzazione iterativi: soluzione per CGLS

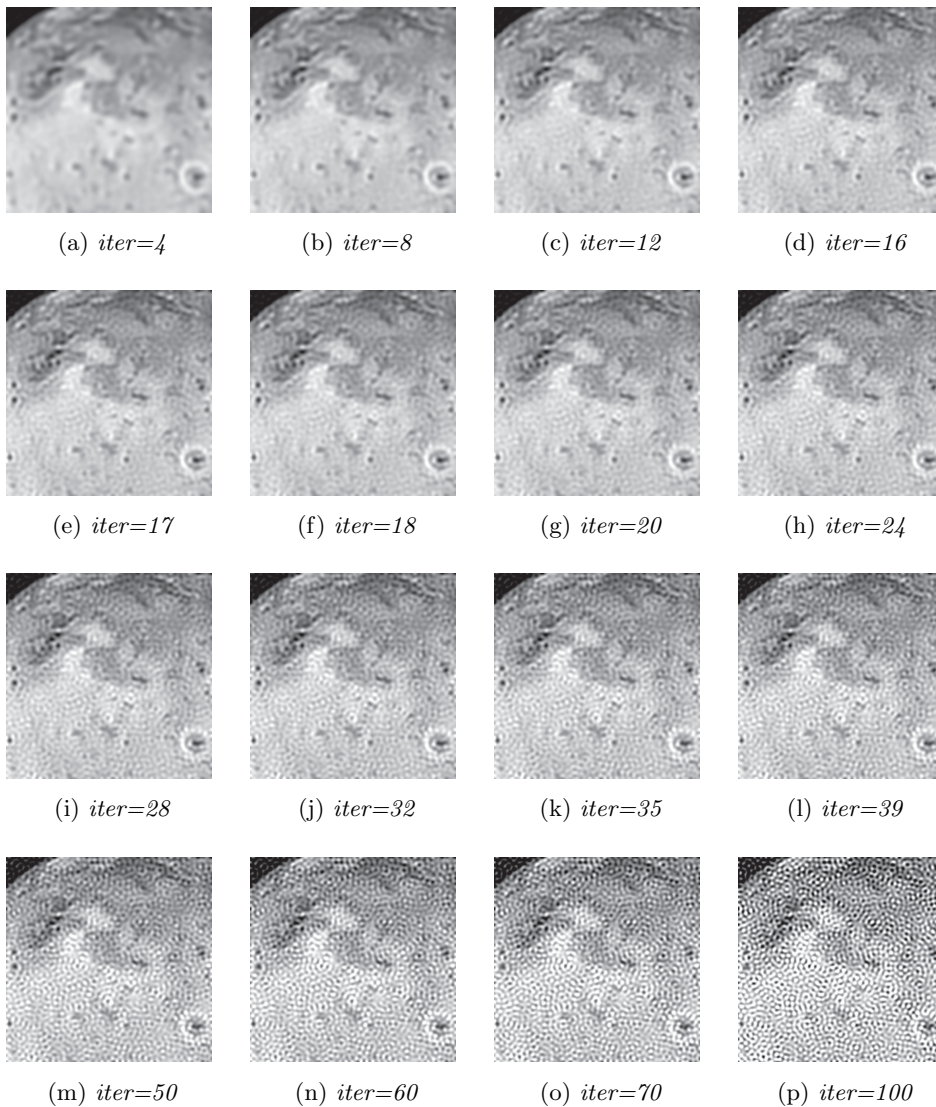
Per l'applicazione del metodo iterativo del CGLS è stata consultata la funzione omonima presente in *Regularization Tools*. Questa routine è adatta per la risoluzione di problemi in dimensione 1 e richiede in input la matrice dei coefficienti A , il termine noto \mathbf{b} e il numero di iterazioni `itermax`. L'output consiste nella matrice X avente `itermax` colonne, ciascuna delle quali contiene la soluzione ottenuta all'iterazione j , con j indice di colonna, per $j = 1, \dots, \text{itermax}$.

Per una necessaria gestione dei costi computazionali e algoritmici, tale funzione è stata modificata nella `X=cglsmodificata(PSF,center,B(:),itermax)`, la quale:

1. non richiede in input la matrice dei coefficienti, che nel nostro caso ha dimensioni $512^2 \times 512^2$;

2. i prodotti matrice-vettore del tipo $A\mathbf{x}$ e $A^t\mathbf{x}$ sono stati ottimizzati sfruttando la particolare struttura della matrice, ovvero, richiamando la funzione `prodottoAvett` già utilizzata in precedenza;
3. restituisce in output la matrice X , avente 512^2 righe e `itermax` colonne; dunque, la j -esima colonna di X contiene l'informazione relativa alla soluzione ottenuta all'iterazione j , per $j = 1, \dots, \text{itermax}$.

Infine, eseguendo un `reshape` delle colonne di X , si possono testare gli effetti del metodo CGLS sulla soluzione al problema considerato.



Osservando quanto ottenuto, si può constatare che dopo circa una ventina di iterazioni si ottengono soluzioni confrontabili con quelle ottenute tramite i metodi di regolarizzazione diretti. Inoltre, al crescere delle iterazioni si ottengono immagini

sempre più affette da rumore, evidenziando, dunque, effetti sulla soluzione simili a quelli prodotti applicando i metodi di filtraggio; in particolare, è evidente la somiglianza con i risultati ottenuti con la SVD troncata per valori crescenti del parametro di troncamento. Questo risultato è concorde col fatto fatto che la soluzione, a ogni iterazione, può essere sempre espressa in termini della SVD come

$$\mathbf{x}^{(k)} = \sum_{i=1}^n f_i^{(k)} \frac{\mathbf{u}_i^t \mathbf{b}}{\sigma_i} \mathbf{v}_i,$$

dove i fattori di filtro $f_i^{(k)}$ tendono a mascherare gli effetti delle componenti ad alte frequenze soltanto per k sufficientemente bassi (cfr. [15]).

Conclusioni

L'analisi condotta in questo elaborato, avvalorata dai test numerici effettuati, ha voluto approfondire il problema del deblurring affrontandolo dal punto di vista matematico, fisico e numerico, al fine di proporre una descrizione completa e che esaltasse le interazioni tra l'analisi numerica e la teoria dei segnali nel comune intento di definire e ricercare modelli e metodi di regolarizzazione per la determinazione di soluzioni fisicamente accettabili per il ripristino di immagini sfocate in bianco e nero. Preliminarmente, è stata posta particolare enfasi sulle proprietà numeriche dei problemi discreti derivanti da problemi di deconvoluzione e sulla struttura matriciale che il problema del deblurring assume sotto alcune ipotesi sull'informazione nei pressi dei bordi dell'immagine; l'importanza di quest'ultimo aspetto è stata esaltata, in modo particolare, nell'implementazione dell'algoritmo CGLS.

Limitandosi a considerare i test effettuati nella parte sperimentale dell'elaborato, si comprende, inoltre, che lo spettro dei metodi numerici applicabili è molto vasto e che le soluzioni da essi indotte, per quanto tutte soddisfacenti, sono, in generale, sempre qualitativamente diverse. Tuttavia, si può riconoscere in tali metodi un aspetto comune, conseguenza delle proprietà matematiche del problema, quello, cioè, legato all'importanza della gestione della quantità e della qualità dell'informazione a disposizione.

È doveroso precisare, infine, che la complessità del fenomeno del deblurring è resa ancora più evidente se lo si considera relativamente al trattamento delle immagini a colori, su cui sono maggiormente focalizzati gli studi delle ricerche odierne e per il cui ripristino sono preferiti metodi di regolarizzazione che generalizzano quelli illustrati in questo lavoro.

Bibliografia

- [1] A. Bjorck, *Numerical Methods for least squares problems*, SIAM (*Society for Industrial and Applied Mathematics*), 1996
- [2] A. Padmanabha Reddy, K.Riyajuddin, *Image deblurring by periodic and Neumann boundary conditions in a wavelet domain*, *Palestine Journal of Mathematics*, Vol. 11(2)(2022) , 617–629, Palestine Polytechnic University-PPU 2022
- [3] A. Quarteroni, R. Sacco, F. Saleri, P. Gervasio, *Matematica Numerica*, 4a edizione, Springer-Verlag Italia 2014
- [4] E. Thiébaud, *Introduction to image reconstruction and inverse problems*, Observatoire de Lyon, <https://cral-perso.univ-lyon1.fr/labo/perso/eric.thiebaut/downloads/documents/cargese-2005-thiebaut.pdf>
- [5] G. C. Barozzi, *Matematica per l'ingegneria dell'informazione*, Zanichelli editore S.p.A, 2004
- [6] G. H. Golub, M. Heath and G. Wahba. *Generalized cross-validation as a method for choosing a good ridge parameter*. *Technometrics*, vol 21: 215-223, 1979
- [7] G. M. Wing, *A primer on Integral Equations of the first kind, The problem of Deconvolution and Unfolding*, SIAM (*Society for Industrial and Applied Mathematics*), Philadelphia, 1991
- [8] G. Rodriguez, *Algoritmi Numerici*, Pitagora Editrice, Bologna, 2008
- [9] G. Rodriguez, D.Theis *An algorithm for estimating the optimal regularization parameter by the L-curve*, *Rendiconti di Matematica*, serie VII Volume 25, Roma (2005), 69-84
- [10] J. M. Varah, *Pitfalls in the Numerical Solution of Linear ill-posed problems*, Computer Science Department, University of British Columbia, Technical Report 81-10, SIAM, 1983

- [11] M. Bertero, P. Boccacci, C. De Mol, *Introduction to Inverse Problems in Imaging*, Second Edition, CRC Press Taylor and Francis Group, 2022
- [12] M. Potmesil, I.Chakravarty, *Synthetic Image Generation with a Lens and Aperture Camera Model*, <https://ftp.bioeng.auckland.ac.nz/jtur044/references/virtual%20environments/p85-potmesil.pdf>
- [13] P. C. Hansen, *Regularization Tools. A Matlab Package for Analysis and Solution of Discrete Ill-Posed Problems*, Marzo 2008
- [14] P. C. Hansen *Deconvolution and regularization with Toeplitz matrices*, Numerical Algorithms 29: 323-378, 2002
- [15] P. C. Hansen, J.G.Nagy, D.P.O'Leary *Deblurring Images, Matrices, Spectra and Filtering*, SIAM (Society for Industrial and Applied Mathematics), Philadelphia, 2006
- [16] P. Diaz de Alba, L. Fermo, F. Pes, G. Rodriguez. *Regularized minimal-norm solution of an overdetermined system of first kind integral equations*, Received: 15 December 2021/Accepted 12 February 2022, Published Online: 05 April 2022, Springer
- [17] R. C. Gonzalez, R.E.Woods, *Digital Image Processing*, Terza Edizione, 2008 by Pearson Education, Inc
- [18] R. M. Gray, Department of Electrical Engineering, Stanford University, USA *Toeplitz and Circulant Matrices: A review*, <https://www-ee.stanford.edu/~gray/toeplitz.pdf>
- [19] S. M. Tan, C.Fox, The University of Auckland, Physics 707, *Inverse Problem, Chapter 1*, https://math.unm.edu/~vageli/courses/Ma375/literature/Inverse_problems_book/chp1.pdf
- [20] W. L. Briggs, V. E. Henson *The DFT, An Owner's Manual for the Discrete Fourier Transform*, SIAM (Society for Industrial and Applied Mathematics), Philadelphia, 1995

Ringraziamenti

Alla fine di questo percorso vorrei porgere alcuni doverosi e sentiti ringraziamenti:

- al professor Giuseppe Rodriguez: per tutto ciò che mi ha insegnato, per aver accettato il ruolo di relatore di questo lavoro e per avermi guidato in maniera esemplare alla sua stesura; per l'eccezionale gentilezza e la costante disponibilità manifestate durante tutto questo percorso;
- ai docenti della facoltà di Matematica, agli amministrativi della Segreteria Studenti e a tutti coloro che sono sempre in prima linea per garantire l'efficienza del servizio e supportare noi studenti nei momenti di dubbio o difficoltà;
- alla mia famiglia, alla mia fidanzata Camilla e a mia nonna Bonaria, che mi hanno sempre sostenuto durante il mio percorso universitario, mostrando orgoglio, comprensione e riconoscenza nei miei confronti. Ma, soprattutto, grazie perchè il coraggio e l'amore che mi avete trasmesso in questi anni sono stati determinanti.
- ai miei colleghi universitari e ai miei amici, per gli scambi di idee, per il confronto e per i momenti felici trascorsi insieme.