



UNIVERSITÀ DEGLI STUDI DI CAGLIARI

FACOLTÀ DI SCIENZE

Corso di Laurea magistrale in Matematica

**RISOLUZIONE DI PROBLEMI INVERSI IN
SPAZI DI BANACH.**

RELATORE:

Prof. Giuseppe Rodriguez

CANDIDATO:

Agus Roberta

Matricola 60/65/65084

Anno accademico 2022/2023

Indice

1	Introduzione	1
1.1	I problemi inversi	1
1.2	Problemi ai minimi quadrati	6
1.3	SVD: Singular Value Decomposition	8
1.4	Metodi di regolarizzazione	10
1.5	Metodi di regolarizzazione iterativi	11
1.6	Teorema di Riesz	12
2	Metodi di regolarizzazione negli spazi di Hilbert	14
2.1	Regolarizzazione: il quadro “classico” negli spazi di Hilbert	14
2.2	TSVD: Truncated Singular Value Decomposition	15
2.3	Regolarizzazione di Tikhonov	16
2.4	Metodi iterativi negli spazi di Hilbert: l’algoritmo di Landweber	18
2.4.1	L’algoritmo di Landweber come procedura iterativa di minimizzazione.	20
2.4.2	L’algoritmo di Landweber come approssimazione iterativa dell’inversa generalizzata $(F^*F)^{-1}F^*$	20
3	Metodi di regolarizzazione negli spazi di Banach	23
3.1	Regolarizzazione: il quadro negli spazi di Banach	23
3.1.1	Alcuni commenti sulla regolarizzazione.	24
3.2	Metodi iterativi negli spazi di Banach: l’algoritmo di Landweber	25
3.2.1	Metodo di Landweber negli spazi L^p	27
4	Esperimenti numerici	30
4.1	Risultati numerici	31
5	Conclusioni	38

Capitolo 1

Introduzione

1.1 I problemi inversi

I problemi inversi sorgono in modo del tutto naturale se si è interessati, ad esempio, a determinare la struttura interna di un sistema fisico dalla misurazione del comportamento del sistema oppure se si è interessati a determinare l'input sconosciuto che dà origine a un segnale di output misurato; a differenza dei problemi diretti in cui l'interesse è nel comportamento del sistema dato l'input o la struttura interna. I problemi inversi lineari sono del tipo:

$$Ax = b$$

dove $b \in Y$ sono i dati osservati, cioè l'effetto. Quello che dobbiamo trovare è un'approssimazione di alcuni parametri del modello $x \in X$, cioè la causa, dove X e Y sono entrambi spazi di Hilbert o di Banach. Nei problemi inversi

$$A : X \longrightarrow Y$$

è spesso un operatore lineare mal posto. L'obiettivo dei problemi inversi, quindi, è determinare l'input o il sistema che dà origine alle misurazioni rumorose dell'output.

Definizione 1.1. Un problema è **ben posto** se:

1. esiste sempre una soluzione,
2. la soluzione è unica,
3. essa dipende in modo continuo dai dati.

I problemi inversi mal posti, violano almeno una delle condizioni presenti nella definizione di problema ben posto. Inoltre, accade spesso che, anche quando sono ben posti, i problemi inversi siano mal condizionati, il che significa che piccoli errori nell' input possono causare errori molto più grandi nell' output. Necessitano quindi di tecniche di regolarizzazione.

Definizione 1.2. Dato un sistema

$$Ax = b$$

in cui il termine noto ha una perturbazione, che può essere un errore di misura o un errore dovuto a calcoli precedenti. In questo caso la soluzione verrà perturbata:

$$A(x + \delta x) = b + \delta b .$$

Il **numero di condizionamento** di questo problema è dato da uno scalare che soddisfa la seguente relazione:

$$\frac{\|\delta x\|}{\|x\|} \leq k(A) \frac{\|\delta b\|}{\|b\|} ,$$

per una norma prefissata.

In sostanza esso dice che l'errore relativo sui dati, e quindi sulle misure, si propaga sulla soluzione di un fattore $k(A)$ che viene detto numero di condizionamento del problema, in questo caso condizionamento relativo perché ci riferiamo a errori relativi. Per un sistema quadrato non singolare è definito come:

$$k(A) = \|A\| \|A^{-1}\| .$$

Il numero di condizionamento può dipendere dai dati o dal problema stesso, quindi non è un numero fissato, in particolare se $k \gg 1$, cioè per k molto grande, si parla di **problema mal condizionato**.

Ricordiamo infine che dobbiamo considerare il fatto che i dati che abbiamo a disposizione nelle applicazioni reali non sono esatti, ma abbiamo a disposizione solo delle loro approssimazioni. Quindi anche del vettore b è disponibile una sua approssimazione, ossia:

$$b = \bar{b} + \epsilon ,$$

dove \bar{b} è il vettore dati esatto.

Vediamo la definizione di spazio di Hilbert che ci serve in seguito nella trattazione.

Definizione 1.3. Uno **spazio di Hilbert** $H = (H, \langle \cdot, \cdot \rangle)$ è uno spazio vettoriale H reale o complesso sul quale è definito un prodotto scalare (o prodotto interno)

$$\begin{aligned} \langle \cdot, \cdot \rangle : H \times H &\longrightarrow \mathbb{R} \\ (x, y) &\longmapsto \langle x, y \rangle \end{aligned}$$

tale che per ogni $x, y, z \in H$ e $\alpha \in \mathbb{R}$ si ha

- $\langle x, x \rangle \geq 0$ e $\langle x, x \rangle = 0 \Leftrightarrow x = 0$, (positività)
- $\langle x, y \rangle = \langle y, x \rangle$, (commutatività)
- $\langle \alpha x, y \rangle = \alpha \langle x, y \rangle$, (omogeneità)
- $\langle x + y, z \rangle = \langle x, z \rangle + \langle y, z \rangle$. (linearità)

Tale spazio deve inoltre essere completo rispetto alla norma indotta dal prodotto scalare:

$$\|x\| := \langle x, x \rangle^{\frac{1}{2}}.$$

Alcuni **esempi** in cui si presentano problemi inversi sono l'acustica, l'astronomia, la tomografia computerizzata, la diffusione elettromagnetica, ottica e restauro di immagini, teoria dello scattering inverso, l'elaborazione dei segnali, ecc... Vediamone alcuni. [4], [2]

ESEMPIO 1: Equazioni integrali di Fredholm di prima specie

Un classico esempio di problema lineare mal posto sono le equazioni integrali di Fredholm di prima specie con nucleo integrabile al quadrato, che possono sempre essere scritte nella forma generale come:

$$\int_0^1 K(s, t) f(t) dt = g(s) , \quad 0 \leq s \leq 1 , \quad (1.1)$$

dove il termine g , solitamente dato in un insieme finito di punti $s = s_i, i = 1, \dots, m$, rappresenta il dato sperimentale, il nucleo K è una funzione nota analiticamente che rappresenta la risposta all'impulso dell'apparecchiatura sperimentale, mentre f è la soluzione cercata, ossia il segnale da recuperare.

Le equazioni integrali di Fredholm del primo tipo modellano problemi fisici che

sorgono in diversi contesti, come elaborazione di immagini, elaborazione di segnali e geofisica.

Da un punto di vista teorico, esse sono trattate in uno spazio di Hilbert che tipicamente coincide con lo spazio delle funzioni integrabili al quadrato. Il corrispondente operatore integrale

$$(Kf)(s) = \int_0^1 k(s,t)f(t) dt$$

è un operatore lineare limitato da uno spazio di Hilbert H_1 in uno spazio di Hilbert H_2 , e una soluzione f di (1.1) esiste solo se g appartiene al range di K , $\mathcal{R}(K) \subset H_2$. Di conseguenza, l'esistenza della soluzione di (1.1) non può essere garantita per qualsiasi membro a destra, ma solo per una classe ristretta di funzioni g . L'unicità della soluzione dipende dalla struttura dello spazio nullo dell'operatore K , ma anche quando è assicurata il problema resta mal posto perché manca la stabilità. In un contesto sperimentale, g è certamente un elemento di $\mathcal{R}(K)$, poiché rappresenta i dati $g(x_i)$ prodotti da un operatore K che riproduce una situazione reale. Questo porta all'equazione integrale con dati discreti:

$$\int_0^1 k_i(t)f(t) dt = b_i, \quad i = 1, \dots, m, \quad (1.2)$$

dove vengono dati m funzionali (o nuclei) k_i su una funzione sconosciuta f , e la (1.2) può essere ottenuta dalla (1.1) con $k_i(t) = K(s_i, t)$ e $b_i = g(s_i)$. Il problema (1.2) è quindi continuo in una sola variabile t . Tuttavia, anche quando $g \in \mathcal{R}(K)$, i valori dei dati in (1.2) sono influenzati da perturbazioni a causa di errori di misura e di arrotondamento, per cui non si può essere sicuri che il membro destro perturbato sia esattamente nel range di K . Inoltre, la soluzione della (1.2) non è unica e non dipende continuamente dai dati. In altre parole, una discretizzazione (1.2) di (1.1) è sempre un problema mal posto.

I prossimi due **esempi** che vediamo sono implementati su matlab in una funzione di Regtools, rispettivamente nel pacchetto deriv2 e shaw.

ESEMPIO 2: Calcolo della derivata seconda (deriv2)

Questo è un classico esempio di problema mal posto. Il nucleo K dell'equazione integrale (1.1) è la funzione di Green della derivata seconda:

$$K(s, t) = \begin{cases} s(t-1), & s < t, \\ t(s-1), & s \geq t, \end{cases}$$

ed entrambi gli intervalli di integrazione sono $[0, 1]$. Il nucleo simmetrico K non è differenziabile nella retta $s = t$. I valori singolari e le funzioni singolari per $i = 1, 2, \dots$ sono dati da:

$$\begin{aligned}\mu_i &= (i\pi)^{-2} , \\ u_i(s) &= \pm\sqrt{2}\sin(i\pi s) , \\ v_i(t) &= \mp\sqrt{2}\sin(i\pi t) .\end{aligned}$$

Poichè i valori singolari sono proporzionali a i^{-2} , il problema è moderatamente mal posto.

Al fine di effettuare esperimenti numerici, fissiamo il membro di destra g e la corrispondente soluzione f in questo modo:

$$g(s) = (s^3 - s)/6 , \quad f(t) = t .$$

La matrice $n \times n$ simmetrica A , il vettore soluzione x^{exact} e il corrispondente membro a destra b^{exact} sono calcolati da K , f e g mediante il metodo Galerkin con basi ortonormali di funzioni definite in $[0, \frac{1}{n}, \frac{2}{n}, \dots, 1]$. Questo porta alle seguenti relazioni per $i = 1, \dots, n$:

$$\begin{aligned}a_{ii} &= h^2 \left(h(i^2 - i + \frac{1}{4}) - (i - \frac{2}{3}) \right) , \\ a_{ij} &= h^2 (j - \frac{1}{2}) \left(h(i - \frac{1}{2}) - 1 \right) , \quad j < i , \\ b_i^{\text{exact}} &= \frac{1}{6} h^{3/2} (i - \frac{1}{2}) \left(\frac{1}{2} h^2 (i^2 - (i - 1)^2) - 1 \right) , \\ x_i^{\text{exact}} &= h^{3/2} (i - \frac{1}{2}) ,\end{aligned}$$

dove $h = \frac{1}{n}$.

ESEMPIO 3: Modello di ricostruzione di immagini in una dimensione (shaw)

Nei problemi di ricostruzione di immagini, si cerca di compensare gli effetti della sfocatura del sistema ottico nell'immagine registrata. Questo problema test utilizza un'equazione integrale di Fredholm di prima specie per modellare una situazione di ricostruzione di immagine in una dimensione. Entrambi gli intervalli di integrazione sono $[-\pi/2, \pi/2]$, e il nucleo K e la soluzione f sono dati da:

$$\begin{aligned}K(s, t) &= (\cos(s) + \cos(t))^2 \left(\frac{\sin(u)}{u} \right)^2 , \\ u &= \pi(\sin(s) + \sin(t)) , \\ f(t) &= a_1 \exp(-c_1(t - t_1)^2) + a_2 \exp(-c_2(t - t_2)^2) .\end{aligned}$$

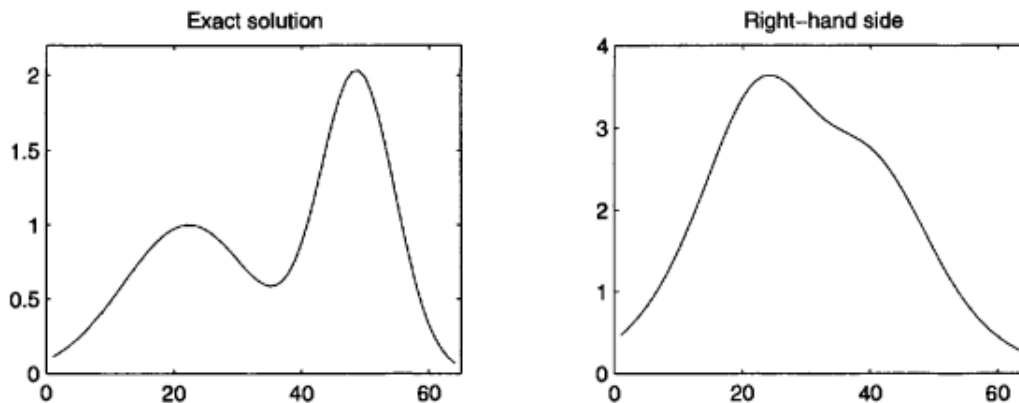


Figura 1.1: La soluzione esatta x^{exact} e il corrispondente membro a destra $b^{\text{exact}} = Ax^{\text{exact}}$ nel problema di ricostruzione di immagine in una dimensione (shaw)

Questo nucleo simmetrico è molto liscio, e i suoi valori singolari sono approssimativamente proporzionali a e^{-2i} ; cioè, il problema è fortemente mal posto. Il nucleo K è la funzione di diffusione dei punti per una fenditura infinitamente lunga di larghezza una lunghezza d'onda, f è l'intensità luminosa della sorgente in funzione dell'angolo t di incidenza, e g è l'intensità luminosa osservata nell'immagine prodotta dall'altro lato della fenditura. I parametri $a_1, a_2, c_1, c_2, t_1, t_2$, sono costanti che determinano l'intensità, l'ampiezza e la posizione delle due sorgenti luminose; qui sono stati utilizzati, $a_1 = 2, a_2 = 1, c_1 = 6, c_2 = 2, t_1 = .8, t_2 = -.5$, dando una f con due "gobbe" distinte. Come mostrato nella figura 1.1, queste due "gobbe" vengono cancellate in g dall'integrazione con K . Il nucleo K e la soluzione f sono discretizzati mediante semplice collocazione con n punti di collocazione, $t_i = (i - 0.5)\pi/n, i = 1, \dots, n$ per produrre la matrice $n \times n$ simmetrica A e il vettore della soluzione x^{exact} . Allora il corrispondente membro di destra discreto viene prodotto come $b^{\text{exact}} = Ax^{\text{exact}}$.

1.2 Problemi ai minimi quadrati

Nelle applicazioni si incontrano spesso delle equazioni del tipo $Ax = b$ con A matrice $m \times n, b \in \mathbb{R}^m$ e $x \in \mathbb{R}^n$ dove, quindi, il numero delle incognite è diverso dal numero delle equazioni. [5] Se supponiamo che A sia a rango pieno potremo avere due casi:

1. $m > n$ il problema potrebbe non avere soluzioni, il sistema si chiama **sovradeterminato**

2. $m < n$ il problema potrebbe avere infinite soluzioni, il sistema si chiama **sottodeterminato**

Questi problemi sono mal posti quindi non hanno una soluzione in senso classico. Fissiamo l'attenzione sul primo caso, ossia quello del sistema **sovradeterminato**. In una situazione ideale, in cui i dati sono esatti, avremo che $m - n$ equazioni sono linearmente dipendenti dalle altre, anche se potremmo non sapere quali. Nel caso in cui i dati sono affetti da errore, invece, può capitare che nessuna delle equazioni sia verificata esattamente, quindi è preferibile utilizzarne un numero sovrabbondante per cercare di ripulire la soluzione dagli errori sperimentali. È quindi ragionevole richiedere che lo scarto quadratico medio (varianza) tra il primo e il secondo membro del sistema sia minimo, passiamo così ad un **problema ai minimi quadrati**. La condizione di scarto quadratico medio si può esprimere nella forma:

$$\min_{x \in \mathbb{R}^n} \|Ax - b\|_2 .$$

Se questa quantità è 0 vuol dire che il sistema ammette una soluzione classica, altrimenti troviamo la soluzione dei minimi quadrati del sistema.

Esistono diverse tecniche per risolvere questi problemi, la più classica è quella che risolve il sistema normale associato al problema.

Volendo minimizzare una quantità non negativa, possiamo considerare il quadrato della norma del residuo $Ax = b$ che in forma matriciale sarà:

$$\begin{aligned} \|Ax - b\|^2 &= (Ax - b)^T(Ax - b) = \\ &= x^T A^T Ax - x^T A^T b - b^T Ax + b^T b = \\ &= x^T A^T Ax - 2x^T A^T b + b^T b . \end{aligned} \quad (1.3)$$

Per minimizzare la norma euclidea del residuo, imponiamo l'annullarsi del gradiente:

$$\frac{1}{2} \nabla(\|Ax - b\|^2) = A^T Ax - A^T b = 0 , \quad (1.4)$$

giungendo così al **sistema normale**:

$$A^T Ax = A^T b . \quad (1.5)$$

Nel caso in cui, come abbiamo supposto all'inizio, A è a rango pieno, la matrice $A^T A \in \mathbb{R}^{n \times n}$ è invertibile e la soluzione del sistema lineare è unica:

$$x_{LS} = (A^T A)^{-1} A^T b. \quad (1.6)$$

Se, invece, A ha rango inferiore a n , tra gli infiniti vettori che soddisfano il sistema, usualmente si assume come soluzione x_{LS} quella dotata di minima norma euclidea.

La matrice

$$A^\dagger = (A^T A)^{-1} A^T$$

viene detta **pseudo-inversa** di A . Essa è un'inversa sinistra di A ma non un'inversa destra.

Nel caso a rango pieno, la matrice $A^T A$ è simmetrica e definita positiva e per risolvere il sistema normale (1.3) si può fare ricorso alla fattorizzazione di Cholesky. Calcolata la fattorizzazione $A^T A = R^T R$, con R matrice $n \times n$ triangolare superiore, il vettore x_{LS} può essere calcolato risolvendo in successione i due sistemi triangolari

$$\begin{cases} R^T y = A^T b \\ R x = y \end{cases} .$$

Il costo computazionale di questo metodo dipende essenzialmente dal calcolo del prodotto $A^T A$ e dalla fattorizzazione di Cholesky, ed è pari a $O\left(\frac{1}{2}n^2(m + \frac{1}{3}n)\right)$ moltiplicazioni. Un possibile svantaggio di questo metodo riguarda la stabilità, dal momento che la matrice $A^T A$ ha un numero di condizionamento pari al quadrato di quello di A .

1.3 SVD: Singular Value Decomposition

Teorema 1.4 (di esistenza).

Sia $A \in \mathbb{R}^{m \times n}$ e supponiamo che $\text{rank}(A) \leq \min(m, n)$, cioè possiamo ammettere che A non abbia rango pieno. Allora :

$\exists U, V$ unitarie con

$$U_{m \times m} \quad \text{t.c.} \quad U^T U = U U^T = I_m \quad ,$$

$$V_{n \times n} \quad \text{t.c.} \quad V^T V = V V^T = I_n \quad ,$$

e inoltre $\exists \Sigma_k = \begin{bmatrix} \sigma_1 & & & \\ & \ddots & & \\ & & \ddots & \\ & & & \sigma_k \end{bmatrix}$ diagonale con $\sigma_i \in \mathbb{R}^+$

convenzionalmente ordinati in ordine decrescente $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_k > 0$ tali che A si possa scrivere come

$$A = U \Sigma V^T$$

dove:

$$\Sigma = \begin{bmatrix} \Sigma_k & 0 \\ 0 & 0 \end{bmatrix} \text{ di dimensione } m \times n ,$$

$$U = [u_1 \quad u_2 \quad \dots \quad u_m] ,$$

$$V = [v_1 \quad v_2 \quad \dots \quad v_n] ,$$

$$\Sigma_k = \text{diag}(\sigma_1 \quad \sigma_2 \quad \dots \quad \sigma_k) .$$

In particolare, gli u_i vengono detti **vettori singolari sinistri**, i v_i **vettori singolari destri** e i σ_i **valori singolari** di A .

Osservazioni

Dal punto di vista geometrico la SVD, fornisce due insiemi di vettori di base ortonormali, ossia le colonne di U e V , in modo tale che la matrice diventi diagonale quando viene trasformata in queste due basi.

In relazione ai problemi discreti mal posti, queste due caratteristiche dell'SVD si trovano molto spesso:

- I valori singolari σ_i decadono gradualmente fino a zero senza particolari divergenze nello spettro. Un aumento delle dimensioni di A aumenterà il numero dei piccoli valori singolari.
- I vettori singolari sinistro e destro u_i e v_i tendono ad avere più cambiamenti di segno nei loro elementi all'aumentare dell'indice i , cioè, come σ_i decresce.

Possiamo quindi scrivere la matrice A in questo modo:

$$A = U\Sigma V^T = \sum_{i=1}^n u_i \sigma_i v_i^T . \quad (1.7)$$

Consideriamo ora la mappatura Ax di un vettore arbitrario x .^[4] Usando l'SVD, otteniamo:

$$x = \sum_{i=1}^n (v_i^T x) v_i \quad (1.8)$$

e

$$Ax = \sum_{i=1}^n \sigma_i (v_i^T x) u_i . \quad (1.9)$$

Queste relazioni mostrano chiaramente che a causa della moltiplicazione con i σ_i , le componenti ad alta frequenza di x sono più smorzate in Ax rispetto alle componenti a bassa frequenza. Inoltre il **problema inverso**, ovvero quello del calcolo di x da $Ax = b$ o da $\min \|Ax - b\|_2$, deve avere l'effetto opposto: amplifica le oscillazioni ad alta frequenza del termine noto b .

Un altro uso della SVD è quello in connessione con i problemi ai minimi quadrati, anche a rango non pieno. Se A è invertibile, allora la sua inversa è data da $A^{-1} = \sum_{i=1}^n v_i \sigma_i^{-1} u_i^T$, e quindi la soluzione di $Ax = b$ è data da:

$$x = \sum_{i=1}^n \sigma_i^{-1} (u_i^T b) v_i .$$

Altrimenti, la pseudoinversa è data da:

$$A^\dagger = \sum_{i=1}^{\text{rank}(A)} v_i \sigma_i^{-1} u_i^T ,$$

e la soluzione dei minimi quadrati x_{LS} del problema ai minimi quadrati $\min \|Ax - b\|_2$ (con minima norma 2 se il $\text{rank}(A) < n$) è dato da:

$$x_{LS} = A^\dagger b = \sum_{i=1}^{\text{rank}(A)} \frac{u_i^T b}{\sigma_i} v_i . \quad (1.10)$$

1.4 Metodi di regolarizzazione

Se il problema è mal condizionato, abbiamo detto che è necessario ricorrere all'uso di un metodo di regolarizzazione, infatti una minima variazione nei dati comporta una grande variazione nella soluzione. Quindi vogliamo trovare una soluzione del problema inverso che non risenta troppo dell'errore e che sia più vicina possibile a quella cercata.

In termini generali, la **regolarizzazione** è l'approssimazione di un problema mal posto con una famiglia di problemi ben posti vicini ad esso. [3] Quello che vogliamo fare è approssimare la soluzione $x^\dagger = A^\dagger b$ di

$$Ax = b$$

per uno specifico termine noto b nella situazione in cui i "dati esatti" \bar{b} non sono noti precisamente, ma di essi abbiamo solo un'approssimazione b^δ , con:

$$\|b^\delta - \bar{b}\| \leq \delta ,$$

dove chiamiamo b^δ i "dati perturbati" e $\delta \in (0, +\infty)$ il "livello di rumore".

Definizione 1.5. Sia $A : X \longrightarrow Y$ un operatore lineare limitato tra spazi di Hilbert X e Y e $\alpha_0 \in (0, +\infty)$. Per ogni $\alpha \in (0, \alpha_0)$, sia

$$R_\alpha : Y \longrightarrow X$$

un operatore continuo, non necessariamente lineare. La famiglia R_α è chiamata **operatore di regolarizzazione** per A^\dagger , se per ogni $b \in \mathcal{D}(A^\dagger)$ esiste una regola di scelta del parametro $\alpha = \alpha(\delta, b^\delta)$ tale che:

$$\lim_{\delta \rightarrow 0} \sup \{ \|R_{\alpha(\delta, b^\delta)} b^\delta - A^\dagger b\| \quad \text{t.c.} \quad b^\delta \in Y, \quad \|b^\delta - b\| \leq \delta \} = 0 . \quad (1.11)$$

Dove

$$\alpha : \mathbb{R}^+ \times Y \longrightarrow (0, \alpha_0)$$

è tale che

$$\lim_{\delta \rightarrow 0} \sup \{ \alpha(\delta, b^\delta) \quad \text{t.c.} \quad b^\delta \in Y, \quad \|b^\delta - b\| \leq \delta \} = 0 . \quad (1.12)$$

Per uno specifico $b \in \mathcal{D}(A^\dagger) := \mathcal{R}(A) + \mathcal{R}(A)^\perp$, la coppia (R_α, α) è chiamata **metodo di regolarizzazione** (convergente) per risolvere $Ax = b$ se la (1.9) e la (1.10) sono verificate. Ricordiamo che $\mathcal{R}(A)$, ossia il **range** di A , è l'insieme di tutti gli elementi del codominio che sono immagine di qualche elemento del dominio tramite A .

1.5 Metodi di regolarizzazione iterativi

Un metodo iterativo funziona come metodo di regolarizzazione se un arresto anticipato delle iterazioni impediscono alle componenti del rumore di interferire (cioè si verifica la semi-convergenza). Il parametro di regolarizzazione è il numero k di iterazioni. Lo schema iterativo di base è:

$$x_{k+1} = x_k - \lambda_k \Phi_A(x_k, b)$$

per la minimizzazione iterativa di funzionali dipendenti dal residuo

$$\frac{1}{r} \|Ax - b\|_Y^r$$

nello spazio di funzioni Y .

L'operatore $\Phi_A : X \times Y \longrightarrow X$ restituisce un valore $\Phi_A(x_k, b)$ che, in sostanza, è un'approssimazione del "gradiente" del funzionale $\frac{1}{r} \|Ax - b\|_Y^r$ nel punto x_k e $\lambda_k > 0$ è la lunghezza del passo. In questo modo, gli schemi iterativi sono tutte diverse generalizzazioni del metodo di discesa di base.

1.6 Teorema di Riesz

Richiamiamo il **teorema di Riesz** che ci servirà quando parliamo di metodi iterativi negli spazi di Hilbert.

Definizione 1.6. Il **duale** di uno spazio è il duale topologico, ossia:

$$X^* = \{\varphi : X \longrightarrow \mathbb{R} \quad \text{tale che } \varphi \text{ è lineare e continuo}\} .$$

Se X è uno spazio normato, anche il suo duale X^* è uno spazio normato con la seguente norma:

$$\|\varphi\|_* = \sup_{\|x\|=1} \varphi(x) .$$

Teorema 1.7 (Riesz). Sia X uno spazio di Hilbert e $\varphi \in X^*$. Allora esiste un'unica $x \in X$ tale che

$$\varphi(y) = \langle x, y \rangle \quad \forall y \in X . \tag{1.13}$$

Inoltre $\|\varphi\|_* = \|x\|$.

Osservazioni

Questo teorema ci dice che tra X e il suo duale topologico X^* c'è un'isometria lineare suriettiva, una mappa che manda tutti gli elementi di X in elementi del duale, che conserva le norme. Questa è un'applicazione lineare e suriettiva perché vale $\forall x \in X^*$. Quindi nel caso degli spazi di Hilbert, X si identifica col suo duale X^* .

Capitolo 2

Metodi di regolarizzazione negli spazi di Hilbert

2.1 Regolarizzazione: il quadro “classico” negli spazi di Hilbert

Sono stati studiati approfonditamente diversi metodi di regolarizzazione per equazioni funzionali mal poste nel contesto degli spazi di Hilbert. Vediamo ora i benefici e gli svantaggi delle tecniche di regolarizzazione in questi spazi.

Vantaggi:

Qualsiasi operatore lineare (o linearizzato) negli spazi di Hilbert, se Hermitiano, può essere scomposto mediante un insieme di autofunzioni utilizzando la teoria spettrale convenzionale. In questo modo, le proprietà di convergenza e regolarizzazione di qualsiasi metodo risolutivo possono essere analizzate considerando il comportamento di ogni singola autofunzione. Nel caso non Hermitiano si procede in modo analogo con la SVD.

Svantaggi:

I metodi di regolarizzazione negli spazi di Hilbert solitamente danno origine a soluzioni lisce (spesso eccessivamente lisce). Nella sfocatura delle immagini, ad esempio, i metodi di regolarizzazione negli spazi di Hilbert non consentono di ottenere una buona localizzazione dei bordi.

2.2 TSVD: Truncated Singular Value Decomposition

Definizione 2.1. L' ϵ -rango numerico di una matrice A , con tolleranza pari a ϵ , è definito da:

$$r_\epsilon = r_\epsilon(A, \epsilon) \equiv \min_{\|E\|_2 \leq \epsilon} \text{rank}(A + E) .$$

In altre parole, l' ϵ -rango di A è uguale al numero di colonne di A che sono linearmente indipendenti per qualsiasi perturbazione di A con norma minore o uguale alla tolleranza ϵ . In termini dei valori singolari di A , l' ϵ -rango numerico, r_ϵ , soddisfa:

$$\sigma_{r_\epsilon} > \epsilon \geq \sigma_{r_\epsilon+1} .$$

In una configurazione ideale, senza perturbazioni ed errori di arrotondamento, l'analisi dei problemi a rango non pieno $Ax = b$ e $\min \|Ax - b\|_2$ è semplice: semplicemente ignora le componenti SVD associate ai valori singolari nulli e calcola la soluzione mediante:

$$x_{LS} = \sum_{i=1}^{\text{rank}(A)} \frac{u_i^T b}{\sigma_i} v_i .$$

In pratica, A non è mai esattamente a rango non pieno, ma piuttosto a rango numerico non pieno; cioè, ha uno o più valori singolari piccoli ma non zero tali che $r_\epsilon < \text{rank}(A) = n$. I piccoli valori singolari inevitabilmente danno luogo a difficoltà. Per capirne il motivo, ricordiamo che la norma di x_{LS} è data da

$$\|x_{LS}\|_2^2 = \sum_{i=1}^n \left(\frac{u_i^T b}{\sigma_i} \right)^2 .$$

Quindi, $\|x_{LS}\|_2$ è molto grande a causa dei piccoli σ_i , a meno che b si trovi pressoché nel range di A , cioè, a meno che gli ultimi $n - r_\epsilon$ coefficienti $u_i^T b$ non soddisfino:

$$|u_i^T b| < \sigma_i, \quad i = r_\epsilon + 1, \dots, n .$$

Ogni volta che sono presenti errori in b , è molto improbabile che questo requisito venga soddisfatto, e la soluzione x_{LS} è quindi dominata dalle ultime $n - r_\epsilon$ componenti della SVD nella (1.8). L'approccio più comune per la regolarizzazione dei problemi a rango numerico non pieno è quello di considerare la matrice A come una rappresentazione perturbata di una matrice matematicamente a rango non pieno, e di sostituire A con una matrice che sia vicina ad essa e matematicamente a rango non pieno. La scelta standard è la matrice A_k di rango k , definita come:

$$A_k \equiv \sum_{i=1}^k u_i \sigma_i v_i^T ;$$

cioè, sostituiamo i piccoli valori singolari diversi da zero $\sigma_{k+1}, \dots, \sigma_n$ con zeri esatti. Tra tutte le matrici Z_k di rango k , la matrice A minimizza entrambe la norma 2 e la norma di Frobenius della differenza $A - Z_k$. È naturale scegliere il rango k di A_k come l' ϵ -rango numerico di A , cioè, $k = r_\epsilon$, perchè $k < r_\epsilon$ porta alla perdita di informazioni associate ai valori singolari grandi, mentre $k > r_\epsilon$ porta a una soluzione con una norma grande e anche una matrice di covarianza con norma grande. Pertanto la regolarizzazione delle soluzioni si ottiene proiettando la matrice mal condizionata A sull'insieme di matrici a rango r_ϵ .

Quando A viene sostituito da A_k , otteniamo un nuovo problema dei minimi quadrati $\min \|A_k x - b\|_2$. La soluzione di minima norma x_k a questo problema, ovvero, l'unica soluzione con norma 2 minima $\|x\|_2$, è data da:

$$x_k = A_k^\dagger b = \sum_{i=1}^k \frac{u_i^T b}{\sigma_i} v_i.$$

La soluzione x_k viene chiamata soluzione SVD troncata, per ovvie ragioni. Il metodo completo è detto SVD troncato (TSVD), e la matrice A_k è detta matrice TSVD. Per riassumere, la soluzione TSVD regolarizzata x_k si ottiene sostituendo prima la matrice mal condizionata A con una matrice A_k di rango k , seguito dal calcolo della soluzione di norma minima dei minimi quadrati di

$$\min \|x\|_2 \quad \text{che dipende da} \quad \min \|A_k x - b\|_2.$$

La norma di x_k è $\|x_k\|_2 = (\sum_{i=1}^k (u_i^T b)^2 \sigma_i^{-2})^{1/2}$, che ovviamente può essere molto più piccola della norma $\|x_{LS}\|_2$ delle soluzioni ai minimi quadrati. Si noti che, come in tutti i problemi di regolarizzazione, raggiungiamo questa riduzione della norma della soluzione consentendo una norma del residuo più grande.

2.3 Regolarizzazione di Tikhonov

Un'alternativa alla TSVD è la regolarizzazione di Tikhonov, introdotta appunto da Tikhonov. Qui si considera il problema regolarizzato:

$$\min_x \|Ax - b\|_2^2 + \lambda^2 \|Dx\|_2^2, \quad (2.1)$$

dove D è una matrice che può assumere diverse forme, tra le più significative abbiamo le seguenti:

$$D = \begin{bmatrix} -1 & 1 & & & & \\ & -1 & 1 & & & \\ & & & \ddots & & \\ & & & & -1 & 1 \\ & & & & & & \end{bmatrix} \quad \text{e} \quad D = \begin{bmatrix} 1 & -2 & 1 & & & \\ & 1 & -2 & 1 & & \\ & & & \ddots & & \\ & & & & 1 & -2 & 1 \end{bmatrix},$$

che approssimano la derivata prima e seconda, rispettivamente.

Il problema (2.1) è anche chiamato problema dei minimi quadrati smorzato, ed equivale al problema dei minimi quadrati

$$\min_x \left\| \begin{pmatrix} \lambda D \\ A \end{pmatrix} z - \begin{pmatrix} 0 \\ b \end{pmatrix} \right\|_2, \quad (2.2)$$

dove la matrice A è stata modificata aggiungendo la matrice λD . Quando $\lambda > 0$ questo problema è sempre a rango di colonna pieno e ha un'unica soluzione.

Se $D = I$ i valori singolari della matrice modificata in (2.2) sono uguali a $\tilde{\sigma}_i = (\sigma_i^2 + \lambda^2)^{1/2}$, $i = 1, \dots, n$. In questo caso la soluzione può essere espressa in termini della SVD come:

$$x(\lambda) = \sum_{i=1}^n \frac{c_i f_i}{\sigma_i} v_i, \quad f_i = \frac{\sigma_i^2}{\sigma_i^2 + \lambda^2}. \quad (2.3)$$

Le quantità f_i sono spesso chiamate **fattori filtro**. Si noti che, se $\lambda \ll \sigma_i$ allora abbiamo che $f_i \approx 1$, mentre se $\lambda \gg \sigma_i$ allora $f_i \approx 0$.

Vediamo come abbiamo ottenuto quest'ultimo risultato espresso nei termini della SVD nel caso in cui $D = I$.

La quantità da minimizzare era $\|Ax - b\|^2 + \lambda^2 \|x\|^2$, quindi andiamo a sostituire la matrice A con la sua decomposizione a valori singolari, ossia $A = U\Sigma V^T$. Attenzione che, affinché questa esistesse, dovevamo avere che $U^T U = U U^T = I_m$ e $V^T V = V V^T = I_n$, ossia le due matrici U e V devono essere unitarie. Facciamo quindi questa sostituzione:

$$\begin{aligned} \|Ax - b\|^2 + \lambda^2 \|x\|^2 &= \|U\Sigma V^T x - b\|^2 + \lambda^2 \|V V^T x\|^2 \\ &= \|\Sigma y - c\|^2 + \lambda^2 \|y\|^2 \end{aligned}$$

dove:

$$y = V^T x \quad \text{e} \quad c = U^T b.$$

Avremo quindi che la x sarà ottenuta da:

$$x = V y.$$

Andiamo ora a risolvere i quadrati delle norme e avremo:

$$y^T \Sigma^T \Sigma y - 2y^T \Sigma^T c + c^T c + \lambda^2 y^T y.$$

Andiamo ora a calcolare il gradiente e poniamo 1/2 di esso pari a zero:

$$\frac{1}{2} \nabla = \Sigma^T \Sigma y - \Sigma^T c + \lambda^2 y = 0,$$

$$(\Sigma^T \Sigma + \lambda^2 I)y = \Sigma^T c ,$$

da cui otteniamo:

$$y_i = \frac{\sigma_i c_i}{\sigma_i^2 + \lambda^2} .$$

La soluzione sarà quindi data da:

$$\begin{aligned} x &= Vy \\ &= \sum_{i=1}^n v_i y_i \\ &= \sum_{i=1}^n \frac{\sigma_i c_i}{\sigma_i^2 + \lambda^2} v_i . \end{aligned}$$

Se, come abbiamo fatto in precedenza, chiamiamo con $f_i = \frac{\sigma_i^2}{\sigma_i^2 + \lambda^2}$, otteniamo esattamente la soluzione come in (2.3). Infatti:

$$x = Vy = \sum_{i=1}^n \frac{\sigma_i c_i}{\sigma_i^2 + \lambda^2} v_i = \sum_{i=1}^n \frac{c_i}{\sigma_i} \frac{\sigma_i^2}{\sigma_i^2 + \lambda^2} v_i = \sum_{i=1}^n \frac{c_i f_i}{\sigma_i} v_i .$$

2.4 Metodi iterativi negli spazi di Hilbert: l'algoritmo di Landweber

Ricordiamo che lo **spazio duale** B^* di uno spazio di Banach B è lo spazio di tutti i funzionali lineari $b^* : B \rightarrow \mathbb{R}$. Se lo spazio di Banach è anche uno spazio di Hilbert, in virtù del Teorema di Riesz, dato un funzionale lineare $b^* \in B^*$, possiamo identificare b^* con un vettore $b \in B$ tale che $b^*(z) = \langle b, z \rangle$, $\forall z \in B$ dove $\langle \cdot, \cdot \rangle$ è il prodotto interno di B . In questo caso, lo spazio duale B^* è isometricamente isomorfo allo spazio B .

Prima di continuare la trattazione e vedere quindi l'algoritmo di Landweber, facciamo vedere che negli spazi di Hilbert l'operatore duale e l'operatore aggiunto sono equivalenti [6]:

Siano X e Y spazi di Banach e X^* e Y^* rispettivamente i loro spazi duali e consideriamo $T : X \rightarrow Y$ un operatore lineare.

Definizione 2.2. $T' : Y^* \rightarrow X^*$ è un **operatore duale** se $\forall x \in X$ e $\forall y' \in Y^*$ si ha:

$$y'(Tx) = (T'y')x. \quad (2.4)$$

Se X e Y sono spazi di Hilbert, per il teorema di Riesz abbiamo che: $\forall y' \in Y^*, \exists y \in Y$ tale che $\forall z \in Y$ si ha:

$$y'(z) = \langle z, y \rangle_Y .$$

Se X e Y sono spazi di Hilbert, consideriamo i seguenti operatori:

$$\begin{aligned} J_Y : y \longmapsto y' & : y'(z) = \langle z, y \rangle_Y \quad \forall z \in Y, \\ J_X : x \longmapsto x' & : x'(t) = \langle t, x \rangle_X \quad \forall t \in X. \end{aligned}$$

Sia ora T' l'operatore duale, ossia $y'(Tx) = (T'y')x$. Allora avremo:

$$\begin{aligned} y'(Tx) &= \langle Tx, J_Y^{-1}y' \rangle_Y , \\ (T'y')x &= \langle x, J_X^{-1}T'y' \rangle_X . \end{aligned}$$

Siccome vale l'uguaglianza (2.4), avremo:

$$\langle Tx, J_Y^{-1}y' \rangle_Y = \langle x, J_X^{-1}T'y' \rangle_X ,$$

cioè:

$$\langle Tx, y \rangle_Y = \langle x, J_X^{-1}T'J_Y y \rangle_X .$$

Dove $J_X^{-1}T'J_Y$ è l' **operatore aggiunto** T^* .

Consideriamo ora il **metodo di Landweber** per l'equazione $Fx = y$ dove $F : X \rightarrow Y$ è lineare ed entrambi gli spazi X e Y sono spazi di Hilbert. Qui consideriamo un dato esatto $y \in \mathcal{R}(F) \oplus \mathcal{R}(F)^\perp \subseteq Y$, essendo $\mathcal{R}(F)$ l'immagine di F (se Y è a dimensione finita, allora $\mathcal{R}(F) \oplus \mathcal{R}(F)^\perp = Y$). Partendo da una soluzione iniziale arbitraria $x_0 \in X$, definiamo le iterazioni come segue

$$x_{k+1} = x_k - \lambda F^*(Fx_k - y), \tag{2.5}$$

dove $\lambda \in (0, \frac{2}{\|F^*F\|})$

Formalmente F^* è l'operatore duale di F , cioè l'operatore $F^* : Y^* \rightarrow X^*$ tale che $(F^*y^*)(x) = y^*(Fx)$, $\forall x \in X$ e $\forall y^* \in Y^*$, dove X^* e Y^* sono gli spazi duali di X e Y . Poiché X è isomorfo a X^* e Y è isomorfo a Y^* , l'operatore duale F^* può essere identificato con l'operatore aggiunto $F^* : Y \rightarrow X$, dove entrambi gli isomorfismi isometrici $\mathcal{I}_Y : Y \rightarrow Y^*$ e $\mathcal{I}_X : X \rightarrow X^*$ sono stati applicati implicitamente. In questo modo, l'iterazione (2.5) è ben definita, poiché l'operatore F^* agisce sul residuo $Fx_k - y^\delta \in Y$, e la sottrazione viene applicata ai due operandi x_k e $\lambda F^*(Fx_k - y)$ di X . Ricordiamo che y^δ è la versione rumorosa dei dati reali y , con livello di rumore δ talvolta noto.

Il metodo di Landweber (2.5) può essere utilizzato in diversi modi. Analizziamo ora due di questi modi.

2.4.1 L'algoritmo di Landweber come procedura iterativa di minimizzazione.

L'algoritmo di Landweber è il metodo più semplice per minimizzare il funzionale convesso¹ $H_2 : X \rightarrow \mathbb{R}$ definito come:

$$H_2(x) = \frac{1}{2} \|Fx - y\|_Y^2, \quad (2.6)$$

dove $\|\cdot\|$ è la norma dello spazio di Hilbert Y , come, ad esempio, $Y = L^2$ o $Y = \ell^2$, con

$$\|f\|_{L^2}^2 = \int_{\mathbb{R}^d} |f|^2 d\mu,$$

oppure nel caso discreto:

$$\|f\|_{\ell^2}^2 = \sum_j |f_j|^2.$$

Il funzionale convesso H_2 è differenziabile e, poiché $y \in \mathcal{R}(F) \oplus \mathcal{R}(F)^\perp$, esiste un insieme non vuoto di punti critici \bar{x} tale che $\nabla H_2(\bar{x}) = 0$ che sono tutti punti minimi di H_2 . Inoltre,

$$-\nabla H_2(x_k) = -F^*(Fx_k - y) \quad (2.7)$$

è esattamente la direzione di discesa più ripida nel punto x_k , quindi l'iterazione di Landweber si sposta da x_k a x_{k+1} proprio lungo la direzione di discesa più ripida. Ancora una volta, con il passo costante $\lambda \in (0, 2/\|F^*F\|)$ si ha $H_2(x_{k+1}) < H_2(x_k)$ a condizione che $H_2(x_k) \neq 0$, e, se $x_0 \in \mathcal{K}(F)^\perp$, garantisce la convergenza a $x^\dagger = F^\dagger y$ della sequenza di iterazioni.

2.4.2 L'algoritmo di Landweber come approssimazione iterativa dell'inversa generalizzata $(F^*F)^{-1}F^*$.

Per induzione è semplice verificare che

$$x_{k+1} = \left(\lambda \sum_{i=0}^k (I - \lambda F^*F)^i F^* y\right) + ((I - \lambda F^*F)^{k+1} x_0). \quad (2.8)$$

Quindi l'algoritmo di Landweber appartiene alla classe dei metodi di Krylov. Infatti, se consideriamo l'ipotesi iniziale fissata $x_0 \equiv 0$, l'iterazione x_{k+1} può essere

¹Si definisce **funzione convessa**, una funzione $f : C \rightarrow \mathbb{R}$, dove C è un sottoinsieme convesso di uno spazio vettoriale X , tale che $f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda)f(y) \quad \forall x, y \in C, \lambda \in [0, 1]$.

scritta come un polinomio di grado k di F^*F applicato ai dati F^*y . Nello specifico, se $Q_k : \mathbb{R} \rightarrow \mathbb{R}$ è il polinomio di grado k definito come

$$Q_k(t) = \lambda \sum_{i=0}^k (1 - \lambda t)^i, \quad (2.9)$$

allora abbiamo

$$x_{k+1} = Q_k(F^*F)F^*y, \quad (2.10)$$

dove $Q_k(F^*F)$ denota un polinomio di F^*F (solo un polinomio matriciale nel caso a dimensione finita). Per calcolo diretto,

$$Q_k(t) = \frac{1 - (1 - \lambda t)^{k+1}}{t} \quad \text{se } t \neq 0,$$

e

$$Q_k(0) = \lambda(k + 1).$$

Quindi, se $t \in (0, 2/\lambda)$ allora $Q_k(t)$ è un'approssimazione del valore t^{-1} , poiché $Q_k(t) \rightarrow t^{-1}$ ($k \rightarrow +\infty$). Inoltre $|1 - \lambda t| < 1$, quindi $(1 - \lambda t)^k \rightarrow 0$, per cui $(I - \lambda F^*F)^k x_0 \rightarrow 0$. Queste due condizioni, insieme all'uniforme limitatezza di $|tQ_k(t)| = |1 - (1 - \lambda t)^{k+1}| \leq 1$, garantiscono che il metodo di Landweber appartiene alla classe degli algoritmi iterativi di regolarizzazione. In questo modo

la sequenza di iterazioni x_k converge alla soluzione ai minimi quadrati $(F^*F)^{-1}F^*y$, ovvero,

$$x_{k+1} = Q_k(F^*F)F^*y + (I - \lambda F^*F)^{k+1}x_0 \rightarrow (F^*F)^{-1}F^*y \quad (2.11)$$

per $k \rightarrow +\infty$, a condizione che esista l'inverso di F^*F e $x_0 \in \mathcal{K}(F)^\perp$.

Più in generale, in un certo senso la k -esima iterazione di Landweber fornisce un'approssimazione regolarizzata $Q_k(F^*F)F^*$ dell'operatore inverso generalizzato $F^\dagger = (F^*F)^{-1}F^*$, che dà x^\dagger quando applicato ai dati $y \in \mathcal{R}(F) \oplus \mathcal{R}(F)^\perp$, cioè, ogni volta che l'equazione ai minimi quadrati $F^*Fx = F^*y$ possiede una soluzione.

Capitolo 3

Metodi di regolarizzazione negli spazi di Banach

3.1 Regolarizzazione: il quadro negli spazi di Banach

Più recentemente, sono stati introdotti e studiati alcuni metodi di regolarizzazione negli spazi di Banach.

Definizione 3.1. Uno **Spazio di Banach** è spazio vettoriale sul campo dei numeri reali o complessi, la cui dimensione può essere infinita, e sul quale è definita una norma tale che ogni successione di Cauchy sia convergente (abbia cioè un limite) a un elemento dello spazio, ossia lo spazio vettoriale è completo.

Una condizione necessaria e sufficiente affinché uno spazio vettoriale normato X sia **completo**, è che tutte le successioni $\{x_n\}_{n=1}^{\infty}$ assolutamente sommabili, cioè tali che:

$$\sum_{n=1}^{\infty} \|x_n\| < \infty, \quad (3.1)$$

siano anche sommabili:

$$\sum_{n=1}^{\infty} x_n < \infty \quad (3.2)$$

e in particolare convergano a un elemento di X .

Vediamo ora i benefici e gli svantaggi delle tecniche di regolarizzazione negli spazi

di Banach.

Vantaggi:

A causa delle proprietà geometriche degli spazi di Banach, questi metodi di regolarizzazione ci permettono di ottenere soluzioni dotate di minore regolarità, che si traducono, ad esempio, in una migliore localizzazione e ripristino delle discontinuità nelle applicazioni di imaging. Un'altra proprietà utile della regolarizzazione negli spazi di Banach è che le soluzioni sono più sparse, cioè, in generale possono essere rappresentate da pochissime componenti, se viene scelta una base opportuna.

Svantaggi:

La teoria spettrale non può essere usata, è richiesta l'analisi convessa.

3.1.1 Alcuni commenti sulla regolarizzazione.

Diversi metodi di regolarizzazione per equazioni funzionali mal poste sono stati formulati come problemi di minimizzazione, dapprima nel contesto degli spazi di Hilbert (ovvero l'approccio classico) e successivamente negli spazi di Banach (ovvero l'approccio più recente).

L'ottimizzazione convessa negli spazi di Banach (in L^1 per un recupero sparso o L^p , $p > 1$ per una regolarizzazione più liscia) aiuta a trarre nuovi algoritmi.

Sostanzialmente possiamo sintetizzare vantaggi e svantaggi nei due spazi, di Hilbert e di Banach, nella tabella seguente:

	Spazi di Hilbert	Spazi di Banach
Vantaggi	Calcolo più semplice (Teoria spettrale, autocomponenti)	Migliore ripristino delle discontinuità; Soluzioni sparse
Svantaggi	Eccessiva uniformità (cattiva localizzazione dei bordi)	Teoricamente complicato (è richiesta l'analisi convessa).

3.2 Metodi iterativi negli spazi di Banach: l'algoritmo di Landweber

L'estensione del metodo di Landweber agli spazi di Banach non è immediata. Innanzitutto, in generale uno spazio di Banach B non è isometricamente isomorfo al suo duale B^* , per cui il termine $F^*(Fx_k - y)$ è privo di significato poiché l'operatore duale $F^* : Y^* \rightarrow X^*$ non può essere applicato a $Fx_k - y \in Y$. Lo stesso avviene per la somma di $x_k \in X$ e $F^*w \in X^*$, qualunque sia $w \in Y^*$. Questo dimostra come la situazione è molto più complicata.

I punti chiave nella generalizzazione dagli spazi di Hilbert a quelli di Banach sono le cosiddette **mappe di dualità**. Una mappa di dualità è una funzione appropriata che ci permette di associare un elemento di uno spazio di Banach B con un elemento del suo duale B^* , ed è utile quando B non è isomorfo a B^* . Siano $b \in B$ e $b^* \in B^*$, in letteratura, per denotare $b^*(b) \in \mathbb{R}$, viene utilizzata la notazione accoppiata $b^*(b) = \langle b^*, b \rangle = \langle b, b^* \rangle$, in questo lavoro di tesi non la utilizziamo per evitare di fare confusione con il prodotto scalare utilizzato negli spazi di Hilbert. Formalmente, per $r > 1$ la mappa di dualità di B , J_r^B è definita come l'operatore multivalore $J_r^B : B \rightarrow 2^{B^*}$, dove 2^{B^*} denota l'insieme di potenze di B^* , tale che

$$J_r^B(b) = \{b^* \in B^* : b^*(b) = \|b\|_B \|b^*\|_{B^*}, \|b^*\|_{B^*} = \|b\|_B^{r-1}\}, \quad \forall b \in B.$$

È interessante notare che, sulla sfera unitaria S_B di B , tutte le mappe di dualità soddisfano la stessa proprietà, cioè la restrizione a S_B di qualsiasi J_r^B per $r > 1$ dà luogo allo stesso operatore $D^B : S_B \rightarrow S_{B^*}$, definito come segue:

$$D^B(b) = \{b^* \in S_{B^*} : b^*(b) = 1\}, \quad \forall b \in S_B.$$

In questo campo si capisce anche che il valore di $r > 1$ non ha un significato essenziale, poiché agisce su tutte le mappe di dualità solo come peso o fattore di scala, in virtù dell'identità:

$$J_r^B(b) = \|b\|_B^{r-2} J_2^B(b), \quad \forall b \in B, b \neq 0.$$

Inoltre, se lo spazio di Banach è anche uno spazio di Hilbert, allora J_2^B è l'isomorfismo isometrico tra B^* e B , e può essere identificato con l'operatore identità in virtù del Teorema di Riesz, cioè $J_2^B(b) = b, \forall b \in B$, ogni volta che B è uno spazio di Hilbert.

La mappa di dualità ha un significato più illustrativo nel contesto della minimizzazione dei funzionali convessi. A questo scopo, sia f un funzionale convesso $f : B \rightarrow \mathbb{R}$. Il sottodifferenziale di f è l'operatore multivalore $\partial f : B \rightarrow 2^{B^*}$ definito come segue:

$$b^* \in \partial f(b) \Leftrightarrow f(c) \geq f(b) + b^*(c - b), \quad \forall c \in B.$$

Un risultato importante è il seguente.

Teorema 3.2 (Asplund). Sia B uno spazio di Banach e sia $r > 1$. Una mappa di dualità J_r^B coincide con il sotto differenziale del funzionale convesso

$$f : B \longrightarrow \mathbb{R} \quad \text{definito come} \quad f(b) = \frac{1}{r} \|b\|_B^r,$$

ovvero

$$J_r^B = \partial f = \partial \left(\frac{1}{r} \|\cdot\|_B^r \right). \quad (3.3)$$

Il **teorema di Asplund** rappresenta la chiave per estendere il metodo di Landweber agli spazi di Banach. Infatti, se consideriamo la generalizzazione del funzionale (2.6) allo spazio di Banach Y e la potenza $r > 1$, otteniamo che il metodo di Landweber corrisponde ora alla procedura di minimizzazione del funzionale convesso $H_r : X \longrightarrow \mathbb{R}$ definito come

$$H_r(x) = \frac{1}{r} \|Fx - y\|_Y^r. \quad (3.4)$$

Infatti, grazie al Teorema di Asplund, il sottodifferenziale di H_r in x può essere calcolato mediante la mappa di dualità $J_r^Y : Y \longrightarrow 2^{Y^*}$ e la regola della catena per la sotto differenziazione delle funzioni composte, abbiamo quindi:

$$\partial H_r(x) = \partial \left(\frac{1}{r} \|Fx - y\|_Y^r \right) = F^* J_r^Y (Fx - y).$$

Osserviamo che ogni elemento dell'insieme $-\partial H_r(x)$ è una direzione di discesa per X^* . Otteniamo così il seguente schema iterativo di Landweber per il problema $Fx = y$ negli spazi di Banach.

L'algoritmo di Landweber negli spazi di Banach

Siano $r, s > 1$ due valori del peso fissati. Sia $x_0 \in X$ un'ipotesi iniziale (spesso nelle applicazioni viene utilizzato il vettore nullo $x_0 = 0 \in X$), e sia $\tilde{x}_0^* = J_s^X(x_0) \in X^*$ un'arbitraria mappa di dualità multivalore. Per $k = 0, 1, 2, \dots$ definiamo:

$$\tilde{x}_{k+1}^* = \tilde{x}_k^* - \lambda F^* J_r^Y (Fx_k - y), \quad (3.5)$$

$$x_{k+1} = J_{s^*}^X(\tilde{x}_{k+1}^*), \quad (3.6)$$

dove s^* è il coniugato di Holder di s , cioè $\frac{1}{s} + \frac{1}{s^*} = 1$.

Qui la mappa di dualità $J_r^X : X \rightarrow 2^{X^*}$ agisce sulle iterazioni $x_k \in X$, per cui la somma in (3.5) è ben definita sullo spazio duale X^* . D'altra parte, la mappa di dualità $J_{s^*}^{X^*} : X^* \rightarrow 2^{X^{**}}$ agisce sulle iterazioni $\tilde{x}_k^* \in X^*$: affinché sia ben definito, è richiesto solo che lo spazio X sia riflessivo, cioè che l'immersione canonica iniettiva e continua j da X in X^{**} sia suriettiva; in questo modo $J_{s^*}^{X^*}$ mappa in X utilizzando implicitamente l'immersione biiettiva j . Osserviamo che qualsiasi mappa di dualità è in generale multivalore e nello schema iterativo di Landweber (3.5)–(3.6) si presuppone implicitamente una scelta arbitraria di un singolo elemento.

L'interpretazione del metodo Landweber come procedura di discesa suggerisce una possibile variante non stazionaria in cui il passo λ in (3.5) dipende da k , secondo la formulazione comunemente usata in letteratura. In questa tesi, per semplicità, non analizziamo questo caso.

3.2.1 Metodo di Landweber negli spazi L^p

Nel caso in cui lo spazio $B = L^p$, possiamo utilizzare per la mappa di dualità l'espressione puntuale

$$J_r^B(b) = \frac{1}{\|b\|_{L^p}^{p-r}} |b|^{p-1} \text{sgn}(b),$$

che preserva la sparsità di b , se presente. Qui “sgn” denota la funzione segno, ovvero $\text{sgn}(b)$ è uguale a $+1$ se $b > 0$, -1 se $b < 0$ e 0 se $b = 0$.

Si parla di soluzioni sparse quando sono costanti a tratti, come ad esempio la soluzione del problema test “tomo” o la soluzione a gradino utilizzata nei problemi test “shaw” e “gravity”, oppure quando esiste una base nella quale pochi coefficienti sono diversi da zero.

Riassumiamo, ora, due principali risultati teorici. [1] Il primo è legato alla forte convergenza delle iterazioni per dati senza rumore. Il secondo ci dice che l'algoritmo è un metodo di regolarizzazione. Si noti che gli spazi di Banach L^p con $1 < p < +\infty$ soddisfano le ipotesi della proposizione seguente.

Proposizione 3.2.1. Sia X uno spazio di Banach liscio e uniformemente convesso¹ (quindi riflessivo) e Y uno spazio di Banach (arbitrario). Sia $y \in \mathcal{R}(\mathcal{F})$ (il range di F) e sia x^\dagger la soluzione di minima norma di $Fx = y$. Se $\lambda_k > 0$ è scelto opportunamente per ogni k , allora la sequenza delle iterazioni x_k definita in

¹Uno spazio di Banach X è detto **uniformemente convesso** se $\forall \epsilon > 0 \exists \delta > 0$ tale che $\forall x, y \in X$ con $\|x\| \leq 1$, $\|y\| \leq 1$ e $\|x - y\| > \epsilon$ risulta $\|\frac{x+y}{2}\| \leq 1 - \delta$.

(3.5)–(3.6) converge fortemente a x^\dagger , cioè

$$\lim_{k \rightarrow \infty} \|x_k - x^\dagger\|_X = 0 .$$

Proposizione 3.2.2. Sia X uno spazio di Banach liscio e uniformemente convesso (quindi riflessivo) e Y uno spazio di Banach (arbitrario). Sia $y \in \mathcal{R}(F)$ e sia $y^\delta \in Y$ un dato rumoroso tali che $\|y^\delta - y\| \leq \delta$.

Se $\lambda_k > 0$ è scelto opportunamente per tutti i k , e viene usato un principio di discrepanza appropriato come regola di stop, allora la sequenza delle iterazioni x_k definita in (3.5)–(3.6) è un metodo di regolarizzazione per $Fx = y$.

Capitolo 4

Esperimenti numerici

In questo capitolo vengono descritti alcuni esperimenti numerici effettuati per verificare l'efficienza del metodo di Landwerber in spazi di Banach e di Hilbert. L'algoritmo verrà analizzato generando diversi sistemi lineari dati da differenti scelte del problema. In particolare sono stati analizzati i seguenti problemi modello estratti dal RegularizationTools di Matlab:

- `shaw(n)`,
- `gravity(n)`,
- `tomo(n)`.

Il problema `shaw`,

Sintassi: $[A, b, x] = \text{shaw}(n)$

ossia il modello di ricostruzione di immagini in una dimensione, è stato analizzato nell'esempio (3) nel capitolo 1. Vediamo ora quali sono gli altri modelli utilizzati.

gravity: Modello di rilevamento gravitazionale 1-D.

Sintassi: $[A, b, x] = \text{gravity}(n, \text{example}, a, b, d)$

Discretizzazione del problema di un modello 1-D per il rilevamento gravitazionale, in cui una distribuzione di massa $f(t)$ si trova alla profondità d , mentre la componente verticale $g(s)$ del campo gravitazionale viene misurata in superficie. Il problema risultante è un'equazione integrale di Fredholm di prima specie del tipo $\int_0^1 K(s, t)f(t) dt = g(s)$, $a \leq s \leq b$, con nucleo

$$K(s, t) = d(d^2 + (s - t)^2)^{-\frac{3}{2}}.$$

Su Matlab sono implementate le seguenti tre soluzioni modello:

1. $f(t) = \sin(\pi t) + 0.5 \sin(2\pi t)$,
2. $f(t) =$ funzione lineare a tratti,
3. $f(t) =$ funzione costante a tratti.

L'esempio 1 è quello di default. Il problema viene discretizzato mediante la regola di quadratura del punto medio con n punti, che porta alla matrice A e al vettore x . Poi il membro destro viene calcolato come $b = Ax$.

L'intervallo di integrazione t è fissato a $[0, 1]$, mentre l'intervallo di integrazione s $[a, b]$ può essere specificato dall'utente. L'intervallo di default è $[0, 1]$, con cui si ottiene una matrice di Toeplitz simmetrica.

Il parametro d è la profondità alla quale si trova il deposito magnetico, e il valore di default è $d = 0.25$. Più grande è la d , più veloce sarà decadimento dei valori singolari. Nelle prove riportate in seguito, sono stati lasciati i valori di default dei parametri a, b, d .

tomo: Problema test della tomografia 2D.

Sintassi: $[A, b, x] = \text{tomo}(N, f)$

Questa funzione crea un semplice problema test di tomografia bidimensionale. Un dominio 2D $[0, N] \times [0, N]$ è diviso in N^2 celle di dimensione unitaria e un totale di raggi pari alla parte intera di $(f \cdot N^2)$ penetra questo dominio in direzioni casuali. Il valore di default è $f = 1$. A ciascuna cella viene assegnato un valore (memorizzato nel vettore x) e per ciascun raggio l'elemento corrispondente nel membro destro b è l'integrale di linea lungo il raggio, cioè:

$$\text{sum}_{\{cells \text{ in ray}\}} x_{\{cell \ j\}} \cdot \text{length}_{\{cell \ j\}},$$

dove $\text{length}_{\{cell \ j\}}$ è la lunghezza del raggio nella j -esima cella.

La matrice A è sparsa e ogni riga (corrispondente a un raggio) contiene il valore $\text{length}_{\{cell \ j\}}$ nella j -esima posizione. Quindi: $b = Ax$. Una volta calcolata una soluzione x_{reg} , può essere visualizzata tramite $\text{imagesc}(\text{reshape}(x_{reg}, N, N))$. La soluzione esatta, $\text{reshape}(x, N, N)$, è identica all'immagine esatta nella funzione sfocatura. Anche in questo caso nelle prove eseguite su Matlab il valore di f utilizzato è quello di default.

4.1 Risultati numerici

Questa sezione riporta i risultati ottenuti utilizzando i problemi test descritti nella sezione precedente, ai fini di verificare l'efficienza dell'algoritmo di Landweber in spazi di Banach e in spazi di Hilbert. Ottenuto $[A, b, sol]$ attraverso i problemi

test sopra descritti, si imposta un valore di λ leggermente più piccolo di $\frac{1}{\text{norm}(A)^2}$. Viene aggiunto quindi del rumore a b attraverso la funzione *addnoise* che prende in ingresso b , un certo ds che viene posto pari a $1e-3$ e il tipo di rumore che nel in questo caso è quello relativo. A questo punto si considera un *noise* con la funzione *randn*($n1, 1$) e viene calcolato il b con il rumore in questo modo: $b = b + ds \cdot \text{norm}(b) \cdot \text{noise}/\text{sqrt}(n1)$. Infine si applica l'algoritmo di Landweber nello spazio di Hilbert o nello spazio di Banach e viene plottato *sol*, ossia la soluzione esatta del problema test e x , cioè l'iterazione k dell'algoritmo di Landweber. L'algoritmo è stato bloccato a mano quando la soluzione ottenuta con il metodo di Landweber approssimava meglio quella esatta e anche provando ad andare avanti con le iterazioni appariva praticamente stabile.

Quello che è stato fatto è confrontare l'efficienza dell'algoritmo di Landweber in spazi di Banach, cambiando il valore di p , e in spazi di Hilbert dove $p = 2$. Questo confronto è stato fatto attraverso un grafico che contenesse tutte e tre le soluzioni, ossia quella esatta, quella ottenuta in spazi di Hilbert e quella ottenuta in spazi di Banach.

Il primo caso considerato, è quello relativo al problema test *shaw*, in cui sono stati considerati $n = 100$ e $\lambda = 0.02$. L'algoritmo di Landweber è stato testato in spazi di Banach con diversi valori del parametro p e si è arrivati alla conclusione che quello che meglio approssimava la soluzione esatta del problema fosse $p = 1.6$.

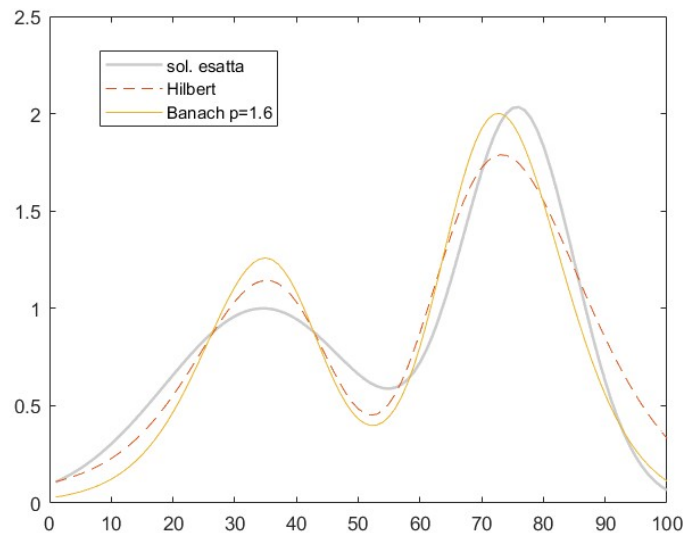


Figura 4.1: Problema test *shaw* in spazi di Hilbert e spazi di Banach con $p = 1.6$ arrestato circa a 1520 iterazioni.

Nella figura 4.1 sono riportati la soluzione esatta, la soluzione negli spazi di Hilbert e la soluzione negli spazi di Banach in sui $p = 1.6$. Come si può osservare nella parte sinistra del grafico il metodo che approssima meglio la soluzione esatta è quello di Hilbert, mentre dall'altra parte e soprattutto il picco viene approssimato meglio dal metodo in spazi di Banach.

Successivamente è stata fatta una variante, ossia è stata costruita una soluzione a gradino sol e il termine noto b non sarà altro che il prodotto tra la matrice A e vettore sol . A questo punto vengono confrontati il metodo di Landweber in spazi di Banach e di Hilbert con questa soluzione a gradino e il risultato è mostrato nel grafico qui sotto.

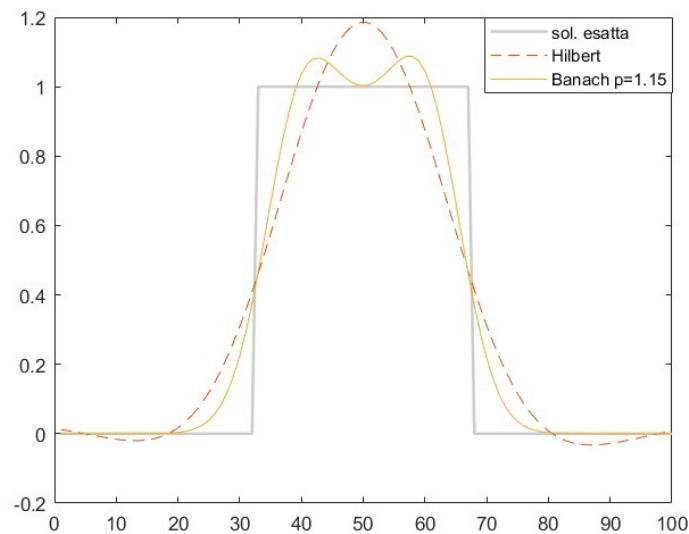


Figura 4.2: Problema test *shaw* in spazi di Hilbert e spazi di Banach con $p = 1.15$ arrestato circa a 4020 iterazioni.

Anche in questo caso sono stati testati diversi valori di p e quello che approssimava meglio la funzione a gradino è $p = 1.15$. In figura 4.2 si può osservare che in questo caso, il metodo che approssima meglio la funzione a gradino è sempre il metodo di Landweber in spazi di Banach.

Il caso successivo è quello relativo al problema test *gravity* con $\lambda = 0.02$ e $n = 100$. Per questo problema, come è stato detto in precedenza, su Matlab sono implementate tre soluzioni modello. Quello che è stato fatto è andare a confrontare le soluzioni del metodo di Landweber in spazi di Hilbert e di Banach con la soluzio-

ne esatta nei primi due casi modello e successivamente con una soluzione gradino costruita su Matlab.

Il primo caso è quello in cui la soluzione modello è data da: $f(t) = \sin(\pi t) + 0.5 \sin(2\pi t)$.

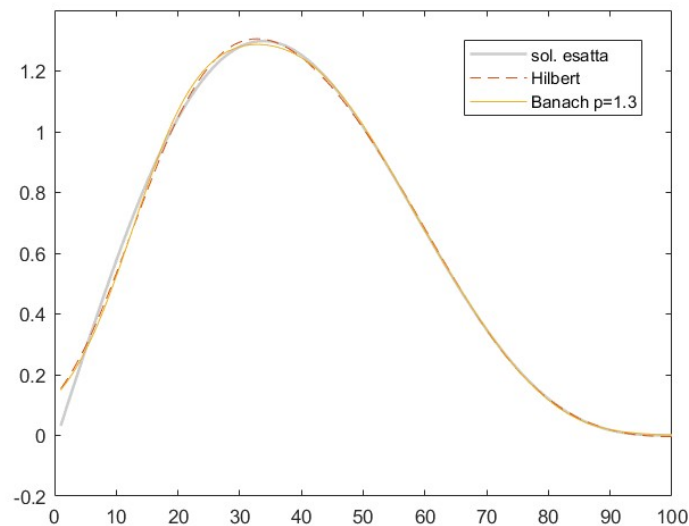


Figura 4.3: Problema test *gravity* in spazi di Hilbert e spazi di Banach con $p = 1.3$ arrestato circa a 1015 iterazioni.

Il valore di p utilizzato negli spazi di Banach in questo caso è stato $p = 1.3$ che è risultato essere il migliore. Come si può osservare in figura 4.3 non ci sono grandi differenze tra il metodo di Landweber in spazi di Hilbert e di Banach. I due metodi coincidono quasi ovunque tranne nel picco che viene approssimato meglio dal metodo negli spazi di Hilbert.

Il secondo caso è quello in cui la soluzione modello è data da: $f(t) =$ funzione lineare a tratti. In questo problema è stato utilizzato p pari a $p = 1.2$ negli spazi di Banach e come si osserva nella figura 4.4 il metodo di Landweber che approssima meglio la soluzione esatta è quello negli spazi di Banach. Questo vale dappertutto, non solo in alcuni punti come accadeva nel problema precedente.

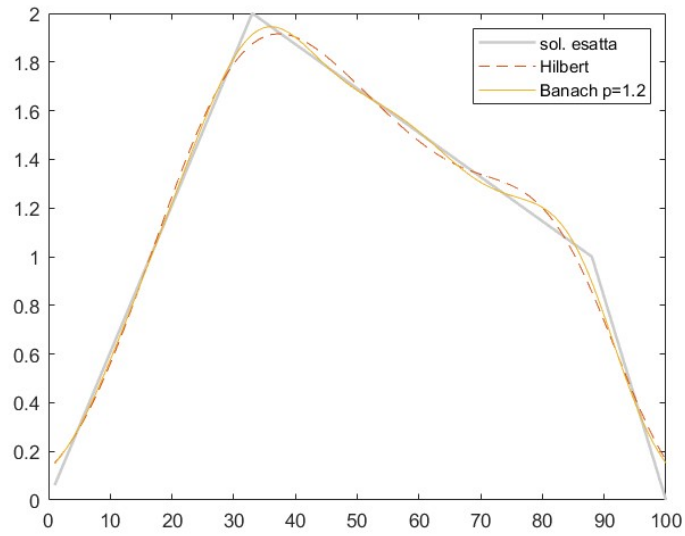


Figura 4.4: Problema test *gravity* in spazi di Hilbert e spazi di Banach con $p = 1.2$ arrestato circa a 2020 iterazioni.

L'ultimo caso analizzato per il problema test *gravity* è quello in cui viene considerata una soluzione a gradino.

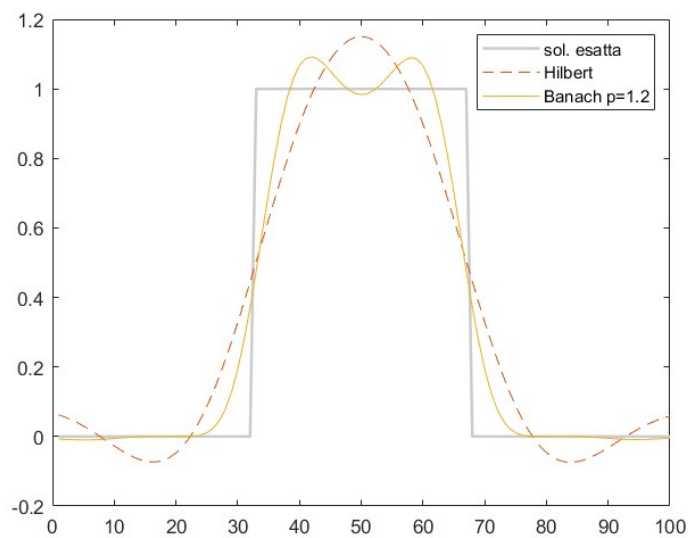
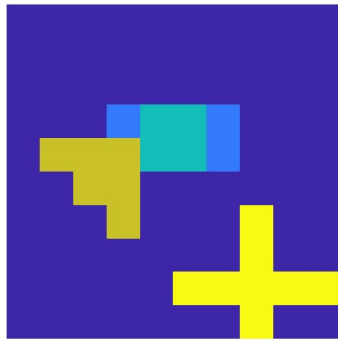


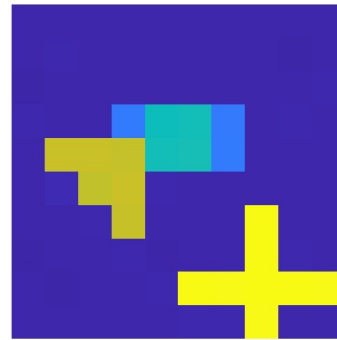
Figura 4.5: Problema test *gravity* in spazi di Hilbert e spazi di Banach con $p = 1.2$ arrestato circa a 1020 iterazioni.

In questo problema test, negli spazi di Banach è stato scelto un p pari a $p = 1.2$ e nella figura 4.5 è possibile osservare i risultati ottenuti. In questo caso, si vede chiaramente che il metodo che approssima meglio la soluzione a gradino è quello utilizzato negli spazi di Banach.

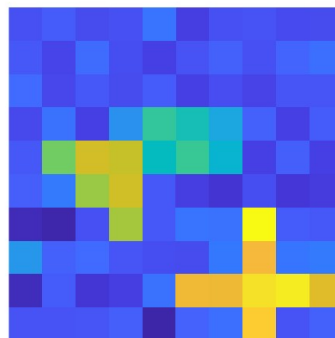
Nel caso del problema test *tomo* sono stati utilizzati un valore di lambda pari a $\lambda = 0.006$ e $n = 10$. In questo problema, ottenuto $[A, b, sol]$, viene aggiunto, come nei casi precedenti, del rumore a b attraverso la funzione `addnoise` e viene plottata la soluzione esatta. Infine si applica l'algoritmo di Landweber in spazi di Hilbert e in spazi di Banach e viene plottata x , ossia l'iterazione dell'algoritmo di Landweber in entrambi i due casi. Quello che bisogna fare è andare a confrontare la soluzione esatta del problema con le soluzioni del metodo di Landweber in spazi di Hilbert e in spazi di Banach, ossia le tre immagini risultanti.



(a) Soluzione esatta



(b) Soluzione in spazi di Banach



(c) Soluzione in spazi di Hilbert

Figura 4.6: Problema test *tomo* arrestato circa a 2000 iterazioni.

Nell'immagine 4.6, abbiamo le tre immagini risultanti; la prima in alto a sinistra è la soluzione esatta, quella a fianco è la soluzione del metodo di Landweber in spazi di Banach in cui è stato considerato $p = 1.22$ e l'immagine sotto è la soluzione del metodo di Landweber in spazi di Hilbert. Come è facilmente osservabile, il metodo che approssima meglio la soluzione esatta è quello in spazi di Banach in cui la soluzione appare più nitida e visibilmente più simile a quella esatta.

Capitolo 5

Conclusioni

La presente tesi si propone di studiare i metodi di risoluzione dei problemi inversi in spazi di Banach. Nel primo capitolo sono stati introdotti i problemi inversi e le tecniche di regolarizzazione necessarie per risolverli. Successivamente, nel secondo capitolo, sono state analizzate le tecniche di regolarizzazione in spazi di Hilbert, in particolare la TSVD, truncated singular value decomposition, la regolarizzazione di Tikhonov ed infine un metodo iterativo che è l'algoritmo di Landweber. Si è poi passati al terzo capitolo in cui è stata studiata l'estensione del metodo di Landweber in spazi di Banach. Quello che è stato fatto nell'ultimo capitolo è testare il metodo di Landweber in spazi L^p e confrontare i risultati in spazi di Hilbert in cui $p = 2$ e in spazi di Banach per diversi valori di p . Quello che ci si aspetta è che il metodo di Landweber sia migliore in spazi di Hilbert quando la soluzione è liscia, infatti come è stato detto in precedenza questo metodo dà luogo, in spazi di Hilbert, a soluzioni anche eccessivamente lisce. Mentre negli spazi di Banach abbiamo un migliore ripristino delle discontinuità. Questo è stato confermato dalle prove numeriche fatte con i tre problemi test.

Bibliografia

- [1] P. Brianzi, F. Di Benedetto, and C. Estatico. Preconditioned iterative regularization in banach spaces. *Computational Optimization and Applications*, 54:263–281, 2013.
- [2] P. Díaz de Alba, L. Fermo, F. Pes, and G. Rodriguez. Regularized minimal-norm solution of an overdetermined system of first kind integral equations. *Numerical Algorithms*, 92(1):471–502, 2023.
- [3] H. W. Engl, M. Hanke, and A. Neubauer. *Regularization of Inverse Problems*, volume 375. Kluwer Academic Publishers, 1996.
- [4] P. C. Hansen. *Rank-Deficient and Discrete Ill-Posed Problems*. SIAM, Philadelphia, 1998.
- [5] G. Rodriguez. *Algoritmi Numerici*. Pitagora Editrice Bologna, 2008.
- [6] K. Yosida. *Functional analysis*, volume 123. Springer, 2012.

Ringraziamenti

Alla fine di questo elaborato, mi sembra doveroso dedicare uno spazio per ringraziare tutte le persone che, con il loro supporto, mi hanno aiutata in questo meraviglioso percorso universitario.

In primis desidero ringraziare il mio relatore, professor Giuseppe Rodriguez, per i suoi preziosi consigli e per la sua disponibilità. Grazie per avermi guidato, con pazienza e gentilezza nella stesura dell'elaborato e per avermi fatto appassionare a questa materia. Un ringraziamento speciale anche a professor Claudio Estatico, per avermi fornito il materiale utile per la stesura di questo elaborato e per la sua disponibilità nell'offrirmi il suo supporto ogni qual volta servisse.

Non posso non ringraziare le due persone che hanno avuto maggiore influenza nel mio percorso educativo, i miei genitori. Mamma e papà, grazie per avermi aiutata a superare i momenti più difficili, senza di voi e i vostri sacrifici non sarei mai potuta arrivare a questo importante traguardo !

Ringrazio infinitamente mia sorella Laura, che mi ha sempre motivata a dare il meglio, che ha condiviso con me gioie e dolori di questo percorso universitario. Grazie per avermi supportata e sopportata in questi anni.

Ringrazio il mio fidanzato Simone per tutto il tempo che mi ha dedicato. Grazie per esserci stato nei momenti belli e per aver avuto sempre le parole giuste per consolarmi nei momenti di sconforto.

Non posso fare a meno di ringraziare i miei nonni, grazie nonna Bonaria mia grande sostenitrice e a voi nonno Raffaele, nonna Letizia e nonno Elia per essere i miei fari anche da lassù.

Grazie a mio zio Antonio, per me come un nonno, sempre presente nella mia vita. Grazie per il tuo costante supporto. Insieme a lui desidero ringraziare tutti i miei zii, zie e cugini che mi stanno accanto sempre.

Grazie a mia figlioccia Alice e suo fratellino Tommaso, per aver sempre fatto il tifo per me e per avermi donato tanti momenti di spensieratezza.

Ringrazio i miei colleghi, in particolare Ilaria, Giorgia, Giuseppe, Nicola e Nour che hanno condiviso con me gioie e fatiche di questi anni trascorsi insieme.

Infine, dedico questa tesi a me stessa, ai miei sacrifici e alla mia tenacia che mi hanno permesso di arrivare fin qui.