

ALMA MATER STUDIORUM · UNIVERSITÀ DI
BOLOGNA

Scuola di Scienze
Laurea Magistrale in Matematica
Curriculum Generale e Applicativo

**Metodo del gradiente coniugato per
problemi ai minimi quadrati non
lineari**

Tesi di Laurea in Analisi Numerica

Relatore:
Chiar.ma Prof.ssa
Germana Landi

Presentata da:
Chiara Gentile

III Sessione
A.a. 2015/2016

Alla mia grande famiglia

Introduzione

Il problema ai minimi quadrati è un problema di ottimizzazione alla base di numerose applicazioni: dalla fisica all'economia, dalla geologia al design (spettroscopia di risonanza magnetica, modellamento di superfici geologiche, allenamento di reti neurali e numerose altre [1]). Più precisamente, data una certa distribuzione di dati sperimentali ed un modello matematico, la risoluzione del relativo problema ai minimi quadrati consente di trovare i parametri del modello che meglio descrivono tali dati. Se i dati in questione sono perturbati da una certa quantità di rumore, la risoluzione del problema ai minimi quadrati potrebbe risentire di tali perturbazioni e produrre una soluzione troppo lontana dalla soluzione del problema non perturbato. Nel caso in cui la soluzione non esiste, o non è unica, o non dipende con continuità dai dati, il problema si dice mal condizionato. Esistono particolari metodi utilizzati in questo contesto, detti di regolarizzazione, che consentono di stimare, a partire da dati con rumore, la soluzione del problema non perturbato. Se il problema ai minimi quadrati è lineare, uno dei metodi più utilizzati è il metodo del gradiente coniugato. Quest'ultimo, nato per l'ottimizzazione di funzioni quadratiche convesse, fornisce buoni risultati anche come metodo di regolarizzazione per problemi ai minimi quadrati lineari mal condizionati. A partire dal 1964 sono state introdotte altre versioni del metodo da poter utilizzare nell'ambito dell'ottimizzazione di funzioni nel caso generale. La prima ad essere proposta è stata quella di Fletcher-Reeves, seguita, nel 1969, dalla variante di Polyak-Polak-Ribière.

L'obiettivo di questo elaborato è analizzare, in analogia al caso lineare, il

possibile utilizzo dei due metodi sopra citati, oltre che nel campo dell'ottimizzazione, anche nel contesto della regolarizzazione di problemi ai minimi quadrati non lineari mal condizionati. Sono stati implementati due problemi inversi mal posti nella forma di equazione integrale di Fredholm, noti nel contesto della gravimetria ([6],[10]). Su entrambi i problemi sono stati testati il metodo di Fletcher-Reeves, il metodo di Polyak-Polak-Ribière e la versione ibrida e poi confrontati con il metodo del gradiente standard e il metodo di Lavenberg-Marquardt. La loro performance è stata analizzata al variare dell'iterato iniziale, della discretizzazione del problema e della quantità di rumore sui dati; in caso di rumore positivo, si è indagato sull'uso del principio della discrepanza come criterio d'arresto. I risultati ottenuti mostrano che i metodi di tipo gradiente coniugato, sebbene siano più lenti nel contesto dell'ottimizzazione rispetto a metodi di tipo Newton, hanno buone proprietà di regolarizzazione: al crescere delle iterazioni i metodi presi in esame si avvicinano alla soluzione esatta maggiormente rispetto agli altri due, con la stessa velocità del metodo di Lavenberg-Marquardt, e se ne allontanano lentamente come il metodo del gradiente; inoltre, arrestando i metodi con il principio della discrepanza si ottengono risultati più soddisfacenti del metodo di Lavenberg-Marquardt. Essi dunque sembrano essere un perfetto compromesso, che sarebbe interessante approfondire per analizzarne l'applicazione a problemi più complessi.

Il mio lavoro si sviluppa in cinque capitoli. Nel primo capitolo viene data la definizione di problema ai minimi quadrati non lineare e viene introdotto il problema del mal condizionamento. Nel secondo capitolo vengono brevemente descritti tre metodi di ottimizzazione non vincolata, attualmente più utilizzati per la risoluzione di problemi ai minimi quadrati non lineari: il metodo di Tikhonov, quello di Gauss-Newton e il metodo di Lavenberg-Marquardt. Successivamente vengono definite le tecniche di regolarizzazione utilizzate in caso di problemi mal condizionati. Il terzo capitolo, invece, è incentrato sul metodo del gradiente coniugato: viene descritto il metodo e riportato uno schema dell'algoritmo classico, ne viene analizzata la conver-

genza e descritte alcune proprietà su cui si fondano i criteri di estensione del metodo al caso generale. Il quarto capitolo è il capitolo centrale dell'elaborato; in esso vengono presentate le due varianti del metodo del gradiente coniugato non lineare prima citate: il metodo di Fletcher-Reeves e il metodo di Polyak-Polak-Ribièr; vengono enunciati i risultati teorici più importanti, confrontati tra loro e introdotta una loro versione ibrida, proposta in [8] da Touati-Ahmed e Storey. Infine, nel quinto capitolo, sono riportati i risultati numerici ottenuti.

Indice

Introduzione	iii
1 Problema ai minimi quadrati non lineare	1
1.1 Definizioni e generalità	1
1.2 Problemi ai minimi quadrati non lineari mal condizionati . . .	3
2 Metodi di risoluzione di problemi ai minimi quadrati non lineari	5
2.1 Metodo di Gauss-Newton	5
2.2 Il metodo di Tikhonov	8
2.3 Il metodo di Lavenberg-Marquardt	8
2.4 Metodi iterativi di regolarizzazione	10
2.4.1 Principio della discrepanza	11
3 Metodo del gradiente coniugato per problemi ai minimi quadrati lineari	13
3.1 Problemi di minimo per funzioni quadratiche convesse	13
3.2 Metodo del gradiente coniugato per funzioni quadratiche convesse	14
3.3 Problemi ai minimi quadrati lineari	16
4 Metodo del gradiente coniugato per problemi ai minimi quadrati non lineari	19
4.1 Metodo del gradiente coniugato per funzioni non convesse . . .	19

4.2	Metodo di Fletcher-Reeves	22
4.2.1	Convergenza del metodo di Fletcher-Reeves	22
4.3	Metodo di Polyak-Polak-Ribière	24
4.3.1	Analogia con il caso quadratico	25
4.3.2	Convergenza del metodo di Polyak-Polak-Ribière	26
4.4	Metodo ibrido FR-PPR	29
5	Risultati numerici	31
5.1	Problema test 1	32
5.2	Problema test 2	44
	Bibliografia	57

Elenco delle figure

5.1	norma del gradiente, test 1, senza rumore	35
5.2	Semiconvergenza dei metodi	37
5.3	Soluzione test 1, rumore 0.1 , 60 punti di discretizzazione . . .	39
5.4	Soluzione test 1, rumore 0.1 , 10 punti di discretizzazione . . .	40
5.5	Soluzione test 1, rumore 0.1 , 120 punti di discretizzazione . .	40
5.6	Soluzione test 1, rumore 0.1 , seed=-2	42
5.7	Soluzione test 1, FR e IBRIDO , $x_0 = \mathbf{ones}(\mathbf{size}(\mathbf{x}_{\mathbf{ex}}))$	43
5.8	Soluzione test 1, PPR e LM, $x_0 = \mathbf{ones}(\mathbf{size}(\mathbf{x}_{\mathbf{ex}}))$	44
5.9	Soluzione test 2	45
5.10	Semiconvergenza test 2, rumore 0.01	47
5.11	Semiconvergenza test 2, PPR e LM, rumore 0.1	48
5.12	Soluzioni test 2, PPR e LM, rumore 0.1	48
5.13	Soluzioni test 2, PPR e LM, rumore 0.1, 120 punti	49
5.14	Soluzioni test 2, IBRIDO e LM, $\mathbf{x}_0 = 0.1 * \mathbf{ones}(\mathbf{size}(\mathbf{x}_{\mathbf{ex}}))$. .	50
5.15	Errore relativo test 2, LM e IBRIDO, rumore 0.01	51
5.16	Soluzioni test 2, IBRIDO e LM, $\mathbf{x}_0 = 0.1 * \mathbf{ones}(\mathbf{size}(\mathbf{x}_{\mathbf{ex}}))$. .	51
5.17	Semiconvergenza test 2, rumore 0.1	52
5.18	Semiconvergenza test 2, rumore 0.1, LM e PPR	53

Capitolo 1

Problema ai minimi quadrati non lineare

In questo capitolo vengono presentati i problemi ai minimi quadrati non lineari e introdotte le notazioni che verranno utilizzate nei capitoli successivi; vengono poi accennate le loro applicazioni a problemi di *datafitting* e introdotte le difficoltà in caso di mal condizionamento.

1.1 Definizioni e generalità

Si definisce *problema ai minimi quadrati non lineare* il seguente problema di ottimizzazione non vincolata

$$\min_x f(x) = \frac{1}{2} \sum_{i=1}^m r_i(x)^2 \quad (1.1)$$

dove le funzioni $r_i : \mathbb{R}^n \rightarrow \mathbb{R}$, dette *residui*, sono funzioni non lineari. Equivalentemente, ponendo

$$r(x) = \begin{pmatrix} r_1(x) \\ r_2(x) \\ \vdots \\ r_m(x) \end{pmatrix}$$

il problema assume la forma

$$\min_x f(x) = \frac{1}{2} \|r(x)\|^2. \quad (1.2)$$

Sia J la matrice jacobiana di r , ossia la matrice $m \times n$ definita come segue

$$J(x) = \begin{pmatrix} \nabla r_1(x) \\ \nabla r_2(x) \\ \vdots \\ \nabla r_m(x) \end{pmatrix} = \left[\frac{\partial r_i}{\partial x_j} \right]_{i=1, \dots, m, j=1, \dots, n}$$

Il gradiente e la matrice hessiana della funzione obiettivo, utilizzando queste notazioni, si scrivono come segue:

$$\nabla f(x) = \sum_{i=1}^m r_i(x) \nabla r_i(x) = J(x)^T r(x)$$

$$\begin{aligned} \nabla^2 f(x) &= \sum_{i=1}^m \nabla r_i(x) \nabla r_i(x)^T + \sum_{i=1}^m r_i(x) \nabla^2 r_i(x) \\ &= J(x)^T J(x) + S(x) \end{aligned}$$

con $S(x) = \sum_{i=1}^m r_i(x) \nabla^2 r_i(x)$.

I metodi numerici di risoluzione di problemi ai minimi quadrati non lineari sono basati in generale sull'assunzione che il termine $S(x)$ sia trascurabile rispetto al termine $J(x)^T J(x)$.

Problemi di questa natura trovano numerose applicazioni, in modo particolare nel *datafitting*, dove spesso i modelli matematici utilizzati per il fit dei dati sono di natura non lineare.

Supponiamo di avere m misure sperimentali

$$(t_1, y_1), (t_2, y_2), \dots, (t_m, y_m)$$

e di voler costruire un modello matematico $M(x, t)$, con $x = (x_1, \dots, x_m)$ parametri del modello, che descriva tali dati. La scelta di $M(x, t)$ dipenderà ovviamente dal tipo di problema considerato.

I parametri x da determinarsi risolvono il problema ai minimi quadrati (1.2) con i residui

$$r_i(x) = y_i - M(x, t_i)$$

1.2 Problemi ai minimi quadrati non lineari mal condizionati

Nelle applicazioni pratiche, le misure sperimentali sono sempre più o meno perturbate da una certa quantità di rumore, dovuta all'azione di agenti esterni quali ad esempio lo strumento di misurazione stesso. È importante dunque andare a studiare l'influenza di tali perturbazioni sulla soluzione del problema ai minimi quadrati perturbato e vedere quanto distante essa sia dalla soluzione reale.

Definizione 1.1. Il problema ai minimi quadrati non lineare (1.2) si dice *mal condizionato* se è sensibile alle perturbazioni sui dati.

La conseguenza del mal condizionamento è l'amplificazione di tali perturbazioni iniziali sulla soluzione.

Nel prossimo capitolo sarà possibile osservare come la matrice jacobiana J giochi un ruolo primario nei metodi numerici di risoluzione del problema e di come dunque, in caso di mal condizionamento, sia necessario adottare metodi opportuni.

Tra i problemi ai minimi quadrati mal condizionati possiamo distinguerne due classi significative: quelli in cui la matrice jacobiana J ha un numero di condizione alto in un intorno di un punto di minimo locale della funzione obiettivo, e quelli che invece nascono dalla discretizzazione di un problema inverso mal posto, caratterizzati da una matrice jacobiana con valori singolari tendenti gradualmente a zero ([1], cap 10). In questo elaborato abbiamo rivolto maggiore attenzione a quest'ultima classe di problemi mal condizionati, su cui sono stati applicati alcuni dei metodi numerici di risoluzione, che saranno descritti successivamente.

Definizione 1.2. Un problema si dice *ben posto* se:

- esiste una soluzione
- la soluzione è unica
- la soluzione dipende con continuità dai dati iniziali

Se una o più di queste condizioni non sono verificate, il problema si dice *mal posto*

Si consideri

$$F(x) = y, \quad (1.3)$$

con $F : \mathbb{R}^n \rightarrow \mathbb{R}$ differenziabile, ottenuta dalla discretizzazione di un problema inverso mal posto.

Come accennato precedentemente, nella realtà si ha sempre a che fare con dati iniziali più o meno perturbati. Si consideri dunque y^δ tali che

$$\|y - y^\delta\|_2 \leq \delta.$$

Si tratta dunque di risolvere il problema inverso

$$F(x) = y^\delta. \quad (1.4)$$

A causa della mal posizione del problema potrebbe non esistere la soluzione x o potrebbe essere troppo distante dalla soluzione di (1.3).

L'approccio classico alla risoluzione di un problema inverso mal posto è la riformulazione del problema in un problema ai minimi quadrati della forma:

$$\min_x \frac{1}{2} \|y^\delta - F(x)\|_2^2 \quad (1.5)$$

il quale, al contrario di (1.4) ha sempre una soluzione, seppure non necessariamente unica, ma che tuttavia risulta essere mal condizionato.

Capitolo 2

Metodi di risoluzione di problemi ai minimi quadrati non lineari

I problemi ai minimi quadrati non lineari sono, in generale, problemi *non convessi* a cui possono essere applicati metodi di ottimizzazione non vincolata; in particolare i più utilizzati sono metodi di ottimizzazione con proprietà regolarizzanti. In questo capitolo verranno analizzati il metodo di *Gauss Newton*, di *Tikhonov* e di *Lavenberg-Marquard*.

2.1 Metodo di Gauss-Newton

Per capire quale sia l'idea alla base del metodo si consideri lo sviluppo di Taylor di secondo grado della funzione obiettivo f nel punto x_k :

$$f(x_k + d) = f(x_k) + d^T \nabla f(x_k) + \frac{1}{2} d^T \nabla^2 f(x_k) d + O(d^3) \quad (2.1)$$

Il metodo di Newton calcola l'iterato successivo $x_{k+1} = x_k + d$ calcolando d che minimizzi la funzione

$$f_*(d) = f(x_k) + d^T \nabla f(x_k) + \frac{1}{2} d^T \nabla^2 f(x_k) d$$

Condizione necessaria affinché d sia un punto di minimo per f_* è che il gradiente di f_* si annulli in d o equivalentemente che valga

$$\nabla^2 f(x_k)d = -\nabla f(x_k) \quad (2.2)$$

$$(J(x_k)^T J(x_k) + S(x_k))d = -J(x_k)^T r(x_k) \quad (2.3)$$

Poichè il calcolo del termine $S(x_k)$ potrebbe essere troppo costoso o le derivate seconde dei residui potrebbero non essere accessibili, il metodo di Gauss Newton si fonda sull'assunzione che il termine $S(x_k)$ sia trascurabile. Questo ovviamente presuppone che l'iterato x_k sia sufficientemente vicino al punto di minimo (equivalentemente che i residui siano sufficientemente piccoli). Viene calcolato dunque il passo d_k come soluzione del sistema

$$J(x_k)^T J(x_k)d = -J(x_k)^T r(x_k) \quad (2.4)$$

Equivalentemente

$$d_k = \operatorname{argmin}_{d \in \mathbb{R}^n} \|J(x_k)d + r_k\|^2 \quad (2.5)$$

Osservazioni:

- il metodo di Gauss-Newton approssima localmente il problema ai minimi quadrati non lineare con un problema ai minimi quadrati lineare.
- Se il rango della matrice $J(x_k)$ è massimo, d_k è sicuramente una direzione di discesa per la f , infatti:

$$\nabla f(x_k)^T d_k = r(x_k)^T J(x_k)d_k = -d_k^T J^T(x_k)J(x_k)d_k = -\|J(x_k)d_k\|^2 \leq 0$$

- A causa del carattere locale dell'approssimazione di Taylor, la lunghezza del passo potrebbe non essere corretta, cioè $f(x_k + d_k) > f(x_k)$. Per ovviare a questo tipo di problema esistono algoritmi specifici per la scelta della lunghezza del passo che consentono di assicurarsi una sufficiente riduzione della funzione ad ogni passo k : un metodo molto

utilizzato, ad esempio, è quello di scegliere un *passo di discesa* α in modo tale che sia soddisfatta la condizione di Armijo:

$$f(x_k + \alpha_k d) < f(x_k) + \alpha_k \nabla f(x_k)^T d \quad (2.6)$$

Eventuali alternative sono i metodo di strategia *trust-region* come il metodo di Lavenberq-Marquardt che vedremo in seguito.

Ricapitolando dunque, un algoritmo generale di questo metodo potrebbe essere il seguente:

Dati:

$x_0 \in \mathbb{R}^n$ e $k = 0$.

While $\nabla f(x_k) \geq tol$

- calcolo di d_k risolvendo

$$\min_{d \in \mathbb{R}^n} \|J(x_k)p + r_k\|^2$$

- calcolo di α_k lungo d_k con una ricerca in linea inesatta.
- $x_{k+1} = x_k + \alpha_k d_k$, $k = k + 1$

End While

Le proprietà di convergenza locale e la velocità di convergenza del metodo dipendono da quanto è grande l'influenza del termine $S(x)$ che è stato omissso nell'approssimazione della matrice hessiana; in particolare, se il problema è consistente, cioè $r(x^*) = 0$, con x^* soluzione esatta, esiste un intorno di x^* in cui la convergenza del metodo è quadratica.

Quest'ultimo è un risultato importante, in quanto, sotto opportune ipotesi, il metodo in questione garantisce le stesse proprietà di convergenza del metodo di Newton con l'evidente vantaggio di non dover calcolare le derivate seconde dei residui. ([3], [1])

2.2 Il metodo di Tikhonov

In caso di problemi ai minimi quadrati mal condizionati, il metodo di Gauss-Newton può essere utilizzato ma risulta poco efficiente, soprattutto se il più piccolo valore singolare della matrice $J(x_k)^T J(x_k)$ tende a zero [3]. In questi casi si preferisce ricorrere a metodi di regolarizzazione; il più famoso è il metodo di Tikhonov.

L'idea alla base del metodo è quella di aggiungere un vincolo sulla norma della soluzione del problema, trasformandolo così in un problema di ottimizzazione vincolata. Utilizzando il metodo dei moltiplicatori di Lagrange si dimostra che risolvere tale problema di minimo vincolato è equivalente a risolvere

$$\min_x \frac{1}{2} (\|r(x)\|_2^2 + \lambda \|x\|_2^2) \quad \lambda \in (0, 1). \quad (2.7)$$

(2.7) è detto **problema di Tikhonov** e il termine aggiunto è detto di *regolarizzazione*, in quanto scegliendo un opportuno valore di λ è possibile compensare il malcondizionamento della matrice jacobiana J . Il comportamento tipico della soluzione al variare della scelta di λ è il seguente:

- $\lambda \rightarrow 1$ il problema (2.7) risulta meglio condizionato ma la soluzione si allontana dalla soluzione del problema originale.
- $\lambda \rightarrow 0$ la soluzione di (2.7) tende alla soluzione di (1.2) ma il problema tende ad essere malcondizionato come il problema originale.

2.3 Il metodo di Lavenberg-Marquardt

Il metodo di Lavenberg-Marquardt nasce propriamente come metodo di regolarizzazione del metodo di Gauss-Newton. Come è stato precedentemente accennato questo metodo utilizza una strategia di tipo *trust-region*: ad ogni passo k cioè, per evitare che il vettore d_k sia troppo grande o abbia una direzione che non assicuri una sufficiente riduzione della funzione obiettivo, si aggiunge un vincolo sulla norma di d_k e invece di risolvere il problema (2.4),

per il calcolo di d_k si risolve il problema

$$\min_d \|J(x_k)d + r(x_k)\|_2^2 + \lambda_k \|d\|_2^2. \quad (2.8)$$

λ_k è il moltiplicatore di Lagrange alla k -esima iterazione; esso influisce sia sulla direzione che sulla lunghezza del passo d_k .

Il passo d_k dunque è soluzione delle equazioni normali:

$$(J(x_k)^T J(x_k) + \lambda_k I)d_k = -J(x_k)^T r(x_k) \quad (2.9)$$

L'aggiunta del termine $\lambda_k I$ serve a diminuire il mal condizionamento della matrice $J(x_k)^T J(x_k)$. Un'altra interpretazione del metodo potrebbe essere quella di vedere il termine $\lambda_k I$ come un'approssimazione del termine $S(x_k)$ che era stato trascurato nel metodo di Gauss-Newton. A tal proposito, il metodo di Lavenberg-Marquardt viene inquadrato nella classe dei metodi di ottimizzazione di secondo grado, o di tipo Newton, che prevedono l'utilizzo della matrice hessiana, o come in questo caso, di una sua approssimazione.

Analogamente al problema di Tikhonov, al variare di λ_k si otterrà una soluzione d_k vicina alla soluzione ottenuta con Gauss-Newton se λ_k è vicino a zero; al crescere di λ_k , invece, il sistema lineare da risolvere risulta meglio condizionato e la soluzione risulta sempre più lontana da quella ottenuta con Gauss Newton.

Per la scelta del parametro λ_k ci sono diverse tecniche: una prima proposta è quella di porre il parametro $\lambda_k = c$ costante ad ogni iterazione k ; altrimenti esistono diverse tecniche per aggiornarlo iterativamente, riportiamo qui di seguito la strategia proposta in [3] che abbiamo implementato nei nostri esperimenti numerici:

Dati:

$$x_0 \in \mathbb{R}^n \quad \text{e} \quad k = 0$$

$$\lambda_0 > 0$$

While $\nabla f(x_k) \geq tol$

- si pone $\lambda_k = \min\{\lambda_0, \|J_k^T r_k\|_2\}$

- risoluzione del sistema lineare (2.9) per il calcolo di d_k
- si determina il passo α_k lungo la direzione d_k con una ricerca in linea inesatta
- $x_{k+1} = x_k + d_k$

End While

Come già accennato precedentemente, i vantaggi nella scelta dell'utilizzo di questo metodo piuttosto che il metodo di Gauss-Newton si riscontrano soprattutto in caso di problemi mal condizionati. ([1], sec 9.3))

2.4 Metodi iterativi di regolarizzazione

In caso di problemi ai minimi quadrati mal condizionati esiste una classe di metodi, detti di **regolarizzazione**, che permettono di stimare, partendo da dati iniziali perturbati da una certa quantità di rumore, la soluzione del problema non perturbato. Il risultato ottenuto con questi metodi dipende da un così detto *parametro di regolarizzazione*; al variare di quest'ultimo la soluzione calcolata potrà essere più o meno vicina alla soluzione esatta. Tale parametro può essere il moltiplicatore di lagrange di un problema di minimo vincolato, come per il metodo di Tikhonov, oppure semplicemente l'iterazione k del metodo. In quest'ultimo caso si parla di **metodi di regolarizzazione iterativi**.

Più precisamente, consideriamo un problema ai minimi quadrati di dati y^δ perturbati da una certa quantità δ di rumore.

Definizione 2.1. Si dice che un metodo di ottimizzazione è un metodo di regolarizzazione iterativo se vale che:

sia $k^*(\delta)$ l'iterazione alla quale viene arrestato il metodo e sia $x_{k^*(\delta)}^\delta$ l'iterato calcolato all'iterazione $k^*(\delta)$, allora

- $x_{k^*(\delta)}^\delta$ è un'approssimazione di una soluzione di (1.3);

- $x_{k^*(\delta)}^\delta$ converge ad una soluzione di (1.3) se δ tende a zero;
- in assenza di rumore, le iterate x_k convergono ad una soluzione di (1.3).

([2], pag 2)

I metodi di regolarizzazione iterativi sono caratterizzati dal fenomeno della **semiconvergenza**. Al variare dell'iterazione k , cioè, l'errore della soluzione calcolata rispetto alla soluzione esatta decresce per alcuni valori del parametro, fino ad arrivare ad un certo valore minimo, per poi ricrescere per valori più grandi del parametro.

2.4.1 Principio della discrepanza

Nei problemi pratici, non conoscendo la soluzione esatta del problema, è difficile scegliere il parametro di regolarizzazione che favorisca il calcolo della soluzione migliore. Tuttavia, sono stati studiati diversi criteri; in questo elaborato prenderemo in considerazione il principio della discrepanza di Morozov [11].

Questo metodo per la scelta del parametro di regolarizzazione k presuppone che sia possibile una stima della norma del rumore, cosa spesso non semplice da ottenere. In particolare, il principio della discrepanza di Morozov sceglie k in modo tale che la corrispondente soluzione regolarizzata x_k soddisfi la seguente relazione

$$\|y - F(x_k)\|_2 = \tau \cdot \delta \quad (2.10)$$

dove τ è una costante positiva e δ è la norma del rumore.

Capitolo 3

Metodo del gradiente coniugato per problemi ai minimi quadrati lineari

In questo capitolo introduciamo brevemente le tecniche di applicazione del metodo del gradiente coniugato a funzioni quadratiche convesse, con particolare riferimento ai problemi ai minimi quadrati lineari. Successivamente vengono analizzate le principali proprietà di tali metodi, utilizzate nel capitolo successivo per l'estensione del metodo al caso non quadratico.

3.1 Problemi di minimo per funzioni quadratiche convesse

Una funzione quadratica su \mathbb{R}^n è una funzione del tipo:

$$q(x) = \frac{1}{2}x^T Qx - c^T x + e \quad (3.1)$$

dove Q è una matrice $n \times n$ simmetrica, $c \in \mathbb{R}^n$ e $e \in \mathbb{R}$.

Proposizione 3.1.1. *Sia q una funzione in \mathbb{R}^n del tipo (3.1).*

- q è convessa se e solo se Q è semidefinita positiva;

- q è strettamente convessa se e solo se Q è definita positiva

Teorema 3.1.2. *Sia q una funzione del tipo (3.1). Allora vale che:*

- q ammette un punto di minimo se e solo se f è convessa (equivalentemente se Q è semidefinita positiva) ed esiste x^* tale che $Qx^* = c$;
- se Q è semidefinita positiva ogni punto x^* tale che $Qx^* = c$ è un punto di minimo globale di f ;
- q ammette un unico punto di minimo globale se e solo se è strettamente convessa (equivalentemente se Q è definita positiva).

Il metodo del gradiente coniugato nasce come metodo di risoluzione di sistemi lineari con matrice dei coefficienti simmetrica e semidefinita positiva, e dunque per quanto detto nel teorema precedente, per la minimizzazione di funzioni quadratiche convesse.

3.2 Metodo del gradiente coniugato per funzioni quadratiche convesse

Si consideri la funzione (3.1), con Q definita positiva. La caratteristica principale del metodo del gradiente coniugato è quella di generare in modo semplice un insieme di direzioni che oltre ad essere linearmente indipendenti sono anche Q -coniugate e ciò consente all'algoritmo di raggiungere il punto di minimo in al più n passi. E' riportato di seguito lo schema classico dell'algoritmo. Per semplificare le notazioni poniamo $g_k = \nabla f(x_k)$.

Dati:

$$x_0 \in \mathbb{R}^n$$

$$\text{Si pone } g_0 = Qx_0 - c, d_0 = -g_0, k = 0.$$

While $\|g_k\|_2 \geq tol$

$$\alpha_k = -\frac{g_k^T d_k}{d_k^T Q d_k} \quad (3.2)$$

$$x_{k+1} = x_k + \alpha_k d_k \quad (3.3)$$

$$g_{k+1} = g_k + \alpha_k Q d_k \quad (3.4)$$

$$\beta_{k+1} = \frac{g_{k+1}^T Q d_k}{d_k^T Q d_k} \quad (3.5)$$

$$d_{k+1} = -g_{k+1} + \beta_{k+1} d_k \quad (3.6)$$

$$k = k + 1$$

End While

Ricavando Qd_k dalla (3.4), è possibile riscrivere la (3.5) nella forma:

$$\beta_{k+1} = \frac{g_{k+1}^T (g_{k+1} - g_k)}{d_k^T (g_{k+1} - g_k)}. \quad (3.7)$$

Si pone, per semplificare le notazioni,

$$y_k = g_{k+1} - g_k$$

e l'espressione finale è

$$\beta_{k+1} = \frac{g_{k+1}^T y_k}{d_k^T y_k}. \quad (3.8)$$

La seguente proposizione mostra che l'algoritmo genera direzioni coniugate e quindi determina il punto di minimo di f in un numero finito di passi:

Proposizione 3.2.1. *L'algoritmo del gradiente coniugato definito dalle (3.2)-(3.8) determina in al più n iterazioni, il punto di minimo della funzione (3.1) con Q definita positiva. In particolare, esiste un intero $m \leq n - 1$ tale che per $i = 1, \dots, m$ si ha*

$$g_i^T g_j = 0, \quad j = 0, 1, \dots, i - 1 \quad (3.9)$$

$$d_i^T Q d_j = 0 \quad j = 0, 1, \dots, i - 1 \quad (3.10)$$

e inoltre risulta $g_{m+1} = 0$.

Vengono elencate adesso una serie di relazioni facili da verificare su cui si fondano i criteri di estensione dell'algoritmo al caso non quadratico.

Osservazione 1. Sotto le ipotesi precedenti valgono le seguenti relazioni:

- moltiplicando scalarmente la (3.4) per d_k si ottiene

$$g_{k+1}^T d_k = 0, \quad (3.11)$$

- moltiplicando scalarmente la (3.6) per g_k si ottiene

$$g_k^T d_k = -g_k^T g_k, \quad (3.12)$$

- sfruttando queste due relazioni si ottiene facilmente

$$d_k^T y_k = -d_k^T g_k = \|g_k\|^2, \quad (3.13)$$

- sfruttando il risultato della proposizione (3.1.1), $g_{k+1}^T g_k = 0$, si ottiene

$$g_{k+1}^T y_k = \|g_{k+1}\|^2. \quad (3.14)$$

La convergenza del metodo in al più n passi è verificata anche nel caso in cui Q sia solo semidefinita positiva, assumendo ovviamente che il sistema $Qx = c$ ammetta almeno una soluzione. Per ulteriori dettagli si veda il capitolo 8 di [3].

3.3 Problemi ai minimi quadrati lineari

Definizione 3.1. Si definisce problema ai minimi quadrati lineare il problema di minimizzazione della seguente funzione:

$$f(x) = \frac{1}{2} \|Ax - b\|^2 = \frac{1}{2} x^T A^T A x - x^T A^T b + \frac{1}{2} b^T b \quad (3.15)$$

dove A è una matrice $m \times n$ e $b \in \mathbb{R}^m$.

La funzione f è una funzione convessa, in quanto la matrice $A^T A$ è sempre semidefinita (o definita) positiva ed è possibile dimostrare che esiste un x^* tale che $A^T A x^* = A^T b$. Pertanto, dal teorema (3.1.2) esiste sempre una soluzione del problema ai minimi quadrati lineari e se la matrice A ha rango massimo tale soluzione è unica e risolve il sistema lineare:

$$A^T A x = A^T b. \quad (3.16)$$

E' riportato qui di seguito uno schema di algoritmo del gradiente coniugato per problemi ai minimi quadrati lineari.

Dati:

$$x_0 \in \mathbb{R}^n$$

$$\text{Si pone } r_0 = b - Ax_0, d_0 = -g_0 = A^T r_0, k = 0.$$

While $\|g_k\|^2 \geq tol$

Si pone

$$q_k = Ad_k \quad (3.17)$$

$$\alpha_k = -\frac{\|g_k\|^2}{\|q_k\|^2} \quad (3.18)$$

$$x_{k+1} = x_k + \alpha_k d_k \quad (3.19)$$

$$r_{k+1} = r_k - \alpha_k q_k \quad (3.20)$$

$$-g_{k+1} = A^T r_{k+1} \quad (3.21)$$

$$\beta_k = \frac{\|g_{k+1}\|^2}{\|g_k\|^2} \quad (3.22)$$

$$d_{k+1} = -g_{k+1} + \beta_{k+1} d_k \quad (3.23)$$

$$k = k + 1$$

End While

Grazie alla condizione (3.21) è possibile interpretare il metodo come un metodo iterativo che agisce direttamente sulla norma del residuo e si ferma nel momento in cui tale norma è abbastanza piccola, cioè l'iterato calcolato è sufficientemente vicino alla soluzione.

Capitolo 4

Metodo del gradiente coniugato per problemi ai minimi quadrati non lineari

Come discusso nel capitolo precedente, i problemi ai minimi quadrati lineari sono problemi *quadratici convessi* che è possibile risolvere con il metodo del gradiente coniugato descritto; i problemi ai minimi quadrati non lineari, invece, non essendo in generale problemi quadratici, nè tantomeno convessi, si risolvono con i metodi di ottimizzazione non vincolata visti nel capitolo 2. Sono state studiate nel corso del tempo diverse versioni del metodo del gradiente coniugato da poter utilizzare anche nel caso di funzioni non quadratiche. In questo capitolo ne vengono introdotte due, viene analizzata la loro convergenza e studiata una loro possibile applicazione a problemi ai minimi quadrati non lineari.

4.1 Metodo del gradiente coniugato per funzioni non convesse

Il metodo descritto nel paragrafo precedente presuppone che la funzione obiettivo sia quadratica. È possibile tuttavia definire delle versioni dell'algo-

ritmo utilizzabili anche in caso di funzioni non quadratiche, che coincidono nel caso quadratico con lo schema definito precedentemente. A tale scopo è importante:

- definire la direzione di ricerca in modo tale che nell'espressione (3.5) non compaia la matrice hessiana Q , la quale potrebbe essere troppo costosa da calcolare nel caso non quadratico;
- sostituire all'espressione (3.2) un'opportuna ricerca unidimensionale della lunghezza del passo α_k .

I vari metodi che sono stati proposti a riguardo differiscono tra loro sia per l'espressione di β_{k+1} , sia per la ricerca unidimensionale adottata per il calcolo di α_k .

Sfruttando le relazioni (3.14) e (3.13) è possibile generare sei formule differenti per β_{k+1} . Ognuna di queste formule associata ad un ricerca in linea, esatta o inesatta, identifica una particolare versione del metodo del gradiente coniugato da utilizzare in caso di funzioni non quadratiche. In questo elaborato verranno prese in considerazione le due formule più note: quella di Fletcher-Reeves e quella di Polyak-Polak-Ribière che è attualmente quella più utilizzata nei codici di calcolo.

- La formula di Fletcher-Reeves si ottiene sostituendo il numeratore della (3.8), sfruttando la relazione (3.14) e lasciando inalterato il denominatore:

$$\beta_{k+1}^{FR} = \frac{\|g_{k+1}\|^2}{\|g_k\|^2}.$$

- La formula di Polyak-Polak-Ribière, invece, si ottiene lasciando inalterato il numeratore di (3.8) e sostituendo il denominatore sfruttando la prima uguaglianza della (3.13):

$$\beta_{k+1}^{PPR} = \frac{g_{k+1}^T y_k}{\|g_k\|^2}.$$

Ulteriori formule si possono generare sommando alle espressioni di β_{k+1} ottenute dei termini, che si annullano nel caso quadratico: se ad esempio β_{k+1}^* è una qualsiasi formula equivalente alla (3.5), allora ogni altra formula del tipo

$$\beta_{k+1} = \beta_{k+1}^* + \xi_k g_{k+1}^T d_k + \zeta_k g_{k+1}^T g_k \quad (4.1)$$

con ξ_k e ζ_k scalari arbitrari sarà ancora una forma equivalente.

Inoltre, poichè, nel caso quadratico

$$\beta_k \geq 0 \quad \text{per ogni } k$$

se β_k^* è una formula equivalente alla (3.5), allora ogni altra formula del tipo

$$\beta_k^+ = \max\{0, \beta_k^*\}$$

è ancora equivalente.

È importante osservare il fatto che le espressioni di β_{k+1} sono equivalenti nel caso quadratico, ma nel caso non quadratico possono definire direzioni di ricerca molto diverse tra loro e dare dunque risultati differenti.

Un'altra questione che è necessario affrontare è la scelta del passo α_k . In caso di funzioni non quadratiche, infatti, non è così semplice trovare un passo che garantisca effettivamente una direzione di discesa; è importante dunque che la tecnica di ricerca unidimensionale sia strettamente correlata alla formula scelta per β_{k+1} , in modo tale da garantire che

$$d_{k+1} = -g_{k+1} + \beta_{k+1} d_k$$

sia effettivamente una direzione di discesa.

Ricordiamo che una condizione sufficiente di discesa è che $g_{k+1}^T d_k < 0$ e dunque:

$$-g_{k+1}^T g_{k+1} + \beta_{k+1} g_{k+1}^T d_k < 0. \quad (4.2)$$

Poichè β_{k+1} , qualunque esso sia, dipende da g_k e $g_{k+1} = g(x_k + \alpha_k d_k)$, il soddisfacimento di tale condizione al passo k dipende esclusivamente dalla scelta del passo α_k . Un'idea, ad esempio, potrebbe essere quella di scegliere α_k in modo tale che uno dei due termini β_{k+1} o $g_{k+1}^T d_k$ sia abbastanza piccolo.

4.2 Metodo di Fletcher-Reeves

Introdotta nel 1964 da Fletcher e Reeves, questo metodo è stata la prima proposta di metodo del gradiente coniugato per funzioni non quadratiche.

Schema del metodo:

Dati:

$$x_0 \in \mathbb{R}^n$$

Si pone $g_0 = g(x_0)$, $d_0 = -g_0$.

While $\|g_k\|^2 \geq tol$

$$\begin{aligned} x_{k+1} &= x_k + \alpha_k d_k \\ \beta_{k+1}^{FR} &= \frac{\|g_{k+1}\|^2}{\|g_k\|^2} \\ d_{k+1} &= -g_{k+1} + \beta_{k+1} d_k \\ k &= k + 1. \end{aligned} \tag{4.3}$$

End While

dove il passo α_k viene calcolato con un'opportuna ricerca unidimensionale.

4.2.1 Convergenza del metodo di Fletcher-Reeves

Il primo risultato di convergenza globale del metodo FR è stato dato da Zoutendijk nel 1970, in particolare egli riuscì a dimostrare la convergenza in caso di utilizzo di una ricerca in linea esatta. A tal proposito enunciamo il seguente teorema.

Teorema 4.2.1. (Convergenza del metodo FR con ricerca esatta)

Sia $f : \mathbb{R}^n \rightarrow \mathbb{R}$ con gradiente Lipschitz-continuo su un insieme aperto convesso \mathcal{D} contenente l'insieme di livello $\mathcal{L}_0 = \{x \in \mathbb{R}^n | f(x) \leq f(x_0)\}$ che supponiamo essere compatto. Sia $\{x_k\}$ una successione generata dal metodo di Fletcher Reeves in cui il passo α_k è determinato nel modo seguente

$$\alpha = \arg \min_{\alpha \geq 0} f(x_k + \alpha d_k).$$

Allora esiste un punto di accumulazione di $\{x_k\}$ che è punto stazionario di f .

([3], sezione 8.4)

Questo risultato, pur essendo rilevante dal punto di vista teorico non è facilmente utilizzabile dal punto di vista pratico, in quanto la ricerca unidimensionale esatta potrebbe essere troppo costosa da effettuare. Tuttavia esso può essere facilmente generalizzato ad un risultato di convergenza globale anche nel caso in cui la ricerca non sia esatta, purchè siano verificate le condizioni di Wolfe.

Definizione 4.1. (Condizioni di Wolfe)

Siano $0 < c_1 < c_2 < 1$, si dice che il passo α soddisfa le condizioni di Wolfe se

$$f(x_k + \alpha d_k) \leq f(x_k) + c_1 \alpha g_k^T d_k \quad (4.4)$$

$$|g_{k+1}^T d_k| \leq c_2 |g_k^T d_k|. \quad (4.5)$$

La prima condizione detta *condizione di Armijo* serve a favorire una sufficiente decrescita della funzione obiettivo, mentre la seconda, detta *condizione della curvatura* è una condizione di sufficiente spostamento, in quanto impedisce che il passo scelto sia troppo piccolo.

Teorema 4.2.2. (Convergenza del metodo FR con ricerca inesatta)

Sia $f : \mathbb{R}^n \rightarrow \mathbb{R}$ con gradiente Lipschitz-continuo su un insieme aperto convesso \mathcal{D} contenente l'insieme di livello \mathcal{L}_0 , che supponiamo essere compatto. Sia $\{x_k\}$ una successione di iterati con $g_k \neq 0$, generata dal metodo di Fletcher-Reeves in cui il passo α_k è determinato in modo tale che, per ogni k , valgano le condizioni di Wolfe, con $c_2 \in (0, 1/2)$. Allora esiste un punto di accumulazione di $\{x_k\}$ che è punto stazionario di f .

([3], sezione 8.4)

Si osserva complessivamente che la convergenza globale del metodo di Fletcher-Reeves, risulta essere di facile verifica; l'unica condizione che si richiede sulla funzione obiettivo è l'ipotesi di lipschitzianità.

Il metodo, tuttavia, risulta essere numericamente spesso inefficiente. Un comportamento tipico, analizzato in [7], è, ad esempio, la tendenza a bloccarsi in regioni 'difficili' dello spazio producendo spostamenti di piccola entità. Un eventuale modifica, che potrebbe essere effettuata all'algoritmo per accelerarne la convergenza, è l'introduzione di quello che viene chiamato '*restart*' sull'antigradiente: se la direzione calcolata dal metodo di FR non produce una discesa sufficiente essa viene sostituita con la direzione dell'antigradiente (direzione di massima discesa).

4.3 Metodo di Polyak-Polak-Ribière

Nel 1969 è stata proposta una nuova versione di gradiente coniugato non lineare: quella di Polak, Ribière e Polyak. Tale versione, nonostante sia ritenuta attualmente una delle più efficienti nel caso non quadratico, presenta maggiori difficoltà nella giustificazione della convergenza globale. Come si vedrà successivamente, non è detto che il metodo converga in caso di ricerca in linea esatta, pertanto è importante trovare tecniche appropriate di ricerca in linea inesatta che garantiscano la convergenza del metodo nel caso generale. Prima di proseguire con alcuni risultati di convergenza del metodo cerchiamo di spiegare in modo euristico perchè esso, nei codici di calcolo, risulta essere più efficiente rispetto al metodo FR.

Consideriamo innanzitutto il coefficiente β del metodo:

$$\beta_k^{PPR} = \frac{g_{k+1}^T y_k}{\|g_k\|^2} \quad \text{con } y_k = g_{k+1} - g_k. \quad (4.6)$$

Si nota da questa formula che nel momento in cui il passo successivo x_{k+1} , calcolato dal metodo, è troppo vicino al passo precedente x_k , si ha che $g_{k+1} \approx g_k$ e $y_k \approx 0$; dunque $\beta_{k+1}^{PPR} \approx 0$ e di conseguenza $d_{k+1} \approx -g_{k+1}$, cioè, la direzione successiva è quella di massima discesa. Si dice in questo caso che il

metodo possiede un restart *automatico* lungo la direzione dell'antigradiente, che salva il metodo dall'effettuare spostamenti troppo piccoli.

4.3.1 Analogia con il caso quadratico

Un'altra possibile motivazione dei migliori risultati del metodo PPR rispetto al metodo FR potrebbe essere il fatto che, in analogia al caso quadratico, il metodo va a calcolare la direzione successiva d_{k+1} in modo tale che sia coniugata alla direzione precedente d_k rispetto a una matrice che rappresenta una media del comportamento dell'hessiana di f . Più precisamente possiamo scrivere, per il teorema del calcolo integrale:

$$\begin{aligned} g_{k+1} - g_k &= \nabla f(x_k + \alpha_k d_k) - \nabla f(x_k) = \\ &= \int_0^1 \frac{d}{dt} \nabla f(x_k + t\alpha_k d_k) dt = \alpha_k B_k d_k. \end{aligned}$$

dove

$$B_k = \int_0^1 \nabla^2 f(x_k + t\alpha_k d_k) dt$$

Si dimostra facilmente a questo punto che le due direzioni sono coniugate rispetto alla matrice B_k , infatti:

$$\begin{aligned} d_{k+1}^T B_k d_k &= (g_{k+1}^T + \beta_{k+1}^{PPR} d_k^T)^T (B_k d_k) = \\ &= -g_{k+1}^T B_k d_k + \frac{g_{k+1}^T (g_{k+1} - g_k)}{d_k^T (g_{k+1} - g_k)} d_k^T B_k d_k = \\ &= -g_{k+1}^T B_k d_k + \frac{g_{k+1}^T (g_{k+1} - g_k)}{(g_{k+1} - g_k)} \frac{(g_{k+1} - g_k)}{\alpha_k} = \\ &= -g_{k+1}^T B_k d_k + \frac{g_{k+1}^T (g_{k+1} - g_k)}{\alpha_k} = 0. \end{aligned}$$

A questo punto dunque basta far in modo che $g_{k+1}^T d_k = 0$, o comunque sia abbastanza piccolo, per ottenere la formula di β_{k+1}^{PPR} e l'analogia è completa.

Questi stessi ragionamenti non possono essere fatti, invece, per il metodo FR. Inoltre, ricordiamo che per il calcolo della formula di β_{k+1}^{FR} , oltre alla condizione $g_{k+1}^T d_k = 0$, si è posta anche la condizione $g_{k+1}^T g_k = 0$, che è più difficile da ottenere nel caso non quadratico.

4.3.2 Convergenza del metodo di Polyak-Polak-Ribière

Come già accennato precedentemente, il metodo PPR se affiancato ad una ricerca in linea esatta, potrebbe non convergere; in [5] sono stati costruiti appositamente esempi in cui effettivamente il metodo non converge. In tal caso bisogna scegliere metodi di ricerca unidimensionale opportuni che garantiscano la convergenza. In particolare, vedremo in questo capitolo metodi di ricerca inesatta attraverso i quali vengono soddisfatte le condizioni sufficienti per la convergenza.

Esiste tuttavia un risultato di convergenza del metodo PPR con ricerca in linea esatta, ma solo nel caso in cui la funzione f sia convessa.

La seguente proposizione, così come le successive, sono tratte da ([3], sezione 8.4).

Proposizione 4.3.1. (Convergenza metodo PPR: caso convesso)

Sia $f : \mathbb{R}^n \rightarrow \mathbb{R}$ due volte continuamente differenziabile su un insieme aperto convesso \mathcal{D} contenente l'insieme di livello \mathcal{L}_0 , che supponiamo essere compatto; siano inoltre $0 < \lambda_1 \leq \lambda_2$ tali che

$$\lambda_1 \|h\|^2 \leq h^T \nabla^2 f(x) h \leq \lambda_2 \|h\|^2, \quad \text{per ogni } x \in \mathcal{L}_0, h \in \mathbb{R}^n. \quad (4.7)$$

Sia $\{x_k\}$ la successione di iterati generata dal metodo PPR, tale che $g_k \neq 0$, in cui il passo α_k è determinato nel modo seguente:

$$\alpha_k = \arg \min_{\alpha \geq 0} f(x_k + \alpha d_k).$$

Allora la successione converge al punto di minimo di f su \mathbb{R}^n .

Per quanto riguarda invece la giustificazione della convergenza con ricerca in linea inesatta, la questione è più complessa. Sotto l'ipotesi che la direzione calcolata dal metodo sia una direzione di discesa, in [4] è dimostrata la convergenza globale del metodo PPR per funzioni strettamente convesse e con una ricerca in linea che soddisfi le condizioni di Wolfe. Tuttavia non è affatto scontato che la direzione calcolata dal metodo sia una direzione di

discesa. A questo proposito sono state studiate diverse modifiche del coefficiente β_k^{PPR} che favoriscono una migliore performance dell'algoritmo. Una di queste, proposta in [5], è la seguente:

$$\beta_k^{PPR+} = \max\{\beta_k^{PPR}, 0\} \quad (4.8)$$

Con questa modifica il metodo risulta essere numericamente più efficiente ed è possibile dimostrare [6] che se la ricerca unidimensionale è tale da soddisfare la condizione di Zoutendijk

$$\sum_{k=0}^m \|g_k\|^2 \cos^2 \theta_k < \infty$$

e la condizione di 'discesa sufficiente'

$$g_k^T d_k \leq -c \|g_k\|^2$$

per qualche $c > 0$, allora esistono dei punti di accumulazione della successione $\{x_k\}$ calcolata dal metodo che sono stazionari di f .

Un'altra possibile strada per migliorare la convergenza del metodo PPR, potrebbe essere quella di apportare modifiche non sul coefficiente β_k ma sull'algoritmo di ricerca in linea inesatta. A tal proposito citiamo alcune condizioni sufficienti per la convergenza globale del metodo, che verranno poi utilizzate per la costruzione di una possibile modifica della ricerca in linea.

Proposizione 4.3.2. (Condizioni di convergenza per il metodo PPR)

Sia $f : \mathbb{R}^n \rightarrow \mathbb{R}$ con gradiente Lipschitz-continuo su un insieme aperto convesso \mathcal{D} contenente l'insieme di livello \mathcal{L}_0 , che supponiamo essere compatto. Sia $\{x_k\}$ la successione di iterati generata dal metodo PPR, tale che $g_k \neq 0$, in cui il passo $\alpha_k > 0$ è determinato in modo tale che, per ogni k , valgano le condizioni seguenti:

$$x_k \in \mathcal{L}_0; \quad (4.9)$$

$$\lim_{k \rightarrow \infty} \frac{|g_k^T d_k|}{\|d_k\|} = 0; \quad (4.10)$$

$$\lim_{k \rightarrow \infty} \|\alpha_k d_k\| = 0. \quad (4.11)$$

Allora si ha

$$\lim_{k \rightarrow \infty} \|g_k\| = 0$$

e dunque esiste un punto di accumulazione per la successione di iterati che è stazionario per f .

L'idea, dunque, è quella di sfruttare queste condizioni sufficienti e produrre una ricerca in linea che permetta di soddisfarle. In particolare, l'algoritmo di Armijio, con opportune modifiche, consente di raggiungere tali risultati.

Vediamo una delle versioni dell'algoritmo modificato con queste proprietà:

Algoritmo AM1 (Armijio modificato)

Dati: $\rho_2 > \rho_1 > 0$, $1 > \eta > 0$, $\sigma \in [0, 1)$, $h \in (0, 1)$.

- Si pone $\tau_k = \frac{|g_k^T d_k|}{\|d_k\|^2}$ e si sceglie $s_k \in [\rho_1 \tau_k, \rho_2 \tau_k]$.
- Si determina $\alpha_k = \max\{h^j s_k, j = 0, 1, \dots\}$ tale che i vettori

$$\begin{aligned} x_{k+1} &= x_k + \alpha_k d_k \\ d_{k+1} &= -g_{k+1} + \beta_{k+1} d_k, \end{aligned}$$

soddisfino le condizioni di Wolfe:

$$f_{k+1} \leq f_k + \eta \alpha_k g_k^T d_k; \quad (4.12)$$

$$g_{k+1}^T d_{k+1} < -\sigma \|g_{k+1}\|^2. \quad (4.13)$$

Enunciamo dunque il teorema di convergenza per il metodo PPR con l'algoritmo di Armijio modificato AM1.

Proposizione 4.3.3. (Convergenza del metodo PPR: ricerca inesatta)

Sia $f : \mathbb{R}^n \rightarrow \mathbb{R}$ con gradiente Lipschitz-continuo su un insieme aperto convesso \mathcal{D} contenente l'insieme di livello \mathcal{L}_0 , che supponiamo compatto. Allora la successione degli iterati, $\{x_k\}$ generata dal metodo PPR, in cui il passo α_k è calcolato con l'algoritmo AM1, ha un punto di accumulazione che è stazionario di f .

4.4 Metodo ibrido FR-PPR

Abbiamo visto dunque che il metodo FR ha buone proprietà di convergenza, al contrario invece del metodo PPR che però risulta essere numericamente più efficiente. A partire da queste osservazioni, Touati-Ahmed e Storey si chiesero se fosse possibile ottenerne un terzo metodo che potesse essere al tempo stesso efficiente e con buone proprietà di convergenza. Quello che proposero, in [8], è un metodo ibrido il cui coefficiente β_k è aggiornato in questo modo:

$$\beta_k = \begin{cases} \beta_k^{PPR} & \text{se } 0 \leq \beta_k^{PPR} \leq \beta_k^{FR} \\ \beta_k^{FR} & \text{altrimenti.} \end{cases} \quad (4.14)$$

Questa formula garantisce che quando l'iterato x_{k+1} è troppo vicino all'iterato x_k , cioè il coefficiente β_k^{PPR} è vicino a zero, viene aggiornato il coefficiente β_k con la formula di Polyak-Polak-Ribière, anziché con la formula di Fletcher-Reeves; in questo modo il metodo evita di compiere spostamenti troppo piccoli, ma allo stesso tempo mantiene le proprietà di convergenza del metodo FR.

Ricordando quanto detto sul coefficiente β_k^{PPR} , in [9] è stato proposto un ulteriore miglioramento da apportare alla formula, in modo tale che il coefficiente calcolato sia sempre positivo:

$$\beta_k = \max\{0, \min\{\beta_k^{FR}, \beta_k^{PPR}\}\}. \quad (4.15)$$

Poiché in generale con questa modifica il metodo risulta più efficiente, nei nostri esperimenti abbiamo implementato l'aggiornamento di β_k come in (4.15)

Capitolo 5

Risultati numerici

Dopo aver introdotto queste tre versioni del metodo del gradiente coniugato a livello teorico come metodi di ottimizzazione non convessa, l'obiettivo è analizzare il comportamento di tali metodi applicati a problemi ai minimi quadrati non lineari mal condizionati. Sono stati dunque implementati il metodo di Fletcher-Reeves, il metodo di Polyak-Polak-Ribière nella versione più efficiente con coefficiente β^{PPR+} e il metodo ibrido; si è scelto di testarli su problemi mal condizionati, per indagare sulle loro proprietà di semiconvergenza e poi confrontarli con il metodo di Lavenberg-Marquardt, precedentemente descritto, e il metodo del gradiente, un metodo di ottimizzazione di primo grado che utilizza come direzione di discesa quella dell'antigradiente ([11]).

Come criterio di arresto per tutti i metodi si è utilizzato

$$\|g_k\| < tol \quad \& \quad k \geq k_{max}$$

con $tol=10^{-8}$ e $k_{max} = 1000$ e, in caso di rumore, è stato implementato il principio della discrepanza, per testarne la sua efficacia all'applicazione di questi metodi.

Per quanto riguarda invece la ricerca in linea è stato implementato l'algoritmo di Armijio con backtracking [12], con condizione di Armijio

$$f(x_k + \alpha_k d_k) \leq f(x_k) + \eta \alpha_k g(x_k)^T \cdot d_k$$

con $\eta = 10^{-1}$ e consentendo fino ad un massimo di 1000 riduzioni del parametro α_k .

5.1 Problema test 1

Il primo problema test considerato, tratto da [6], è un problema inverso noto nell'ambito della gravimetria, nella forma di equazione integrale di Fredholm del primo ordine non lineare

$$F(x)(t) = \int_a^b k(t, s, x(s)) ds = y(t), \quad t \in [e, f] \quad (5.1)$$

con kernel

$$k(t, s, x(s)) = \ln \frac{(t-s)^2 + H^2}{(t-s)^2 + (x(s) - H)^2} \quad (5.2)$$

e dati $y(t)$.

Il kernel k è definito sull'insieme $\Pi = \{[e, f] \times [a, b] \times R\}$, $k(t, s, x(s)) \in C^1(\Pi)$ e $\frac{\partial k}{\partial x}(t, s, x(s)) = \frac{2(H-x(s))}{(t-s)^2 + (x(s)-H)^2}$.

Per risolvere numericamente (5.1), per prima cosa discretizziamo l'intervallo $[e, f]$ scegliendo $e = t_0 < t_1 < \dots < t_m = f$; dopodichè scegliamo funzioni linearmente indipendenti $\{\phi_j\}_{j=1}^n \subset X = H_0^1(a, b)$ e consideriamo l'approssimazione di $x(s)$, $\hat{x}(s) = \sum_{j=1}^n x_j \phi_j(s)$, con $x = (x_1, x_2, \dots, x_n)^T \in \mathbb{R}^n$.

Scegliamo come ϕ_j le funzioni lineari standard:

$$\phi_j(s) = \begin{cases} \frac{s-s_{j-1}}{h}, & \text{if } s \in [s_{j-1}, s_j], \\ \frac{s_{j+1}-s}{h}, & \text{if } s \in [s_j, s_{j+1}] \\ 0, & \text{altrimenti} \end{cases}$$

e $s_j = a + hj$, $h = \frac{b-a}{n}$, $j = 1, 2, \dots, n$.

Si osserva che

$$\phi_j(s_k) = \begin{cases} 1 & \text{se } k = j \\ 0 & \text{altrimenti} \end{cases} \implies \hat{x}(s_k) = \sum_{j=0}^n x_j \phi_j(s_k) = x_k \quad (5.3)$$

Discretizzando in questo modo, F diventa un operatore $F : \mathbb{R}^n \rightarrow \mathbb{R}^m$,

$$[F_{mn}(x)]_i = \int_a^b k(t_i, s, \hat{x}(s)) ds, \quad 1 \leq i \leq m. \quad (5.4)$$

Analogamente l'operatore derivata F' diventa una matrice $m \times n$:

$$[F'_{mn}(x)]_{ji} = \int_a^b \frac{\partial}{\partial x_j} k(t_i, s, \hat{x}(s)) ds \quad 1 \leq i \leq m, \quad 1 \leq j \leq n; \quad (5.5)$$

più precisamente,

$$[F'_{mn}(x)]_{ji} = \int_a^b \frac{\partial k}{\partial x_j}(t_i, s, \hat{x}(s)) \frac{\partial \hat{x}}{\partial x_j}(s) ds = \int_a^b \frac{\partial k}{\partial x_j}(t_i, s, \hat{x}(s)) \phi_j(s) ds$$

Ricordando la proprietà (5.3) delle funzioni ϕ_j e risolvendo gli integrali con la formula dei trapezi [13] si ottiene:

$$\begin{aligned} [F'_{mn}(x)]_{ji} &= 0.5 \cdot h \cdot \sum_{k=1}^{n-1} \left(\frac{\partial k}{\partial x_k}(t_i, s_k, x_k) \phi_j(s_k) + \frac{\partial k}{\partial x_{k+1}}(t_i, s_{k+1}, x_{k+1}) \phi_j(s_{k+1}) \right) = \\ &= 0.5 \cdot h \cdot 2 \cdot \left(\frac{\partial k}{\partial x_j}(t_i, s_j, x_j) \right). \end{aligned}$$

Dunque,

$$[F'_{mn}(x)]_{ji} = \begin{cases} 0.5 \cdot h \cdot \frac{\partial k}{\partial x_1}(t_i, s_1, x_1) & \text{se } j = 1, \\ h \cdot \frac{\partial k}{\partial x_j}(t_i, s_1, x_j) & \text{se } j = 2, \dots, n-1 \\ 0.5 \cdot h \cdot \frac{\partial k}{\partial x_n}(t_i, s_n, x_n) & \text{se } j = n. \end{cases}$$

Il problema continuo mal posto (5.1) può essere discretizzato in un problema ai minimi quadrati mal condizionato, del tipo:

$$\min_{x \in \mathbb{R}^n} \frac{1}{2} \|F(x) - y\|^2 \quad (5.6)$$

Scegliamo $[a, b] = [e, f] = [0, 1]$, $H = 0.1$ e diversi valori di m e n per la discretizzazione degli intervalli. Il problema ammette due soluzioni continue: $x_{ex}(s) = 1.3s(s-1)$ e $x_{ex}(s) = 1.3s(1-s) + 0.2$; anch'esse vengono discretizzate valutandole nei punti s_i dell'intervallo $[a, b]$.

Per quanto riguarda invece il termine a destra di (5.1), per simulare problemi pratici può essere perturbato in questo modo:

$$y^\delta = y_{ex} + \text{noise} \cdot \text{rand}(\text{size}(y)),$$

dove

$$y_{ex} = \int_a^b k(t, s, x_{ex}(s)) ds$$

viene calcolato utilizzando la formula dei trapezi; ovviamente per la costruzione della funzione obiettivo è necessario fare una scelta tra le due funzioni soluzioni del problema, nei nostri esperimenti abbiamo fatto riferimento alla funzione $x_{ex}(s) = 1.3s(1-s) + 0.2$. Il rumore viene fatto variare scegliendo $\text{noise} \in \{0.1, 0.01, 0.001\}$ e il comando `rand` di Matlab genera un vettore di numeri casuali; in particolare, utilizzando il comando `rand('seed', n)`, è possibile provare al variare di `n` diverse versioni di vettori casuali.

Il problema (5.6) è dunque un problema ai minimi quadrati non lineari che può essere risolto con i metodi precedentemente descritti; la funzione obiettivo è

$$f(x) = \frac{1}{2} \|F(x) - y\|^2$$

e il gradiente della funzione obiettivo è

$$g(x) = F'(x)^T (F(x) - y).$$

Riportiamo di seguito alcuni dei risultati più significativi al variare dell'iterato iniziale x_0 , del rumore, e dei valori di m e n .

- $x_0 = 0.2 * \text{ones}(\text{size}(x_{ex}))$,

$m = n = 60$, `noise=0`

	Errore relativo	norma del gradiente	iter
CG FR	1.45e-03	1.76e-06	1000
CG PPR	1.66-03	3.15-07	1000
CG IBRIDO	1.49e-03	3.9e-07	1000
Lavenberg-Marquardt	1.05e-03	5.37e-09	28
Gradiente	3.54e-03	2.67e-05	1000

Visualizziamo l'andamento della norma del gradiente nei vari metodi.

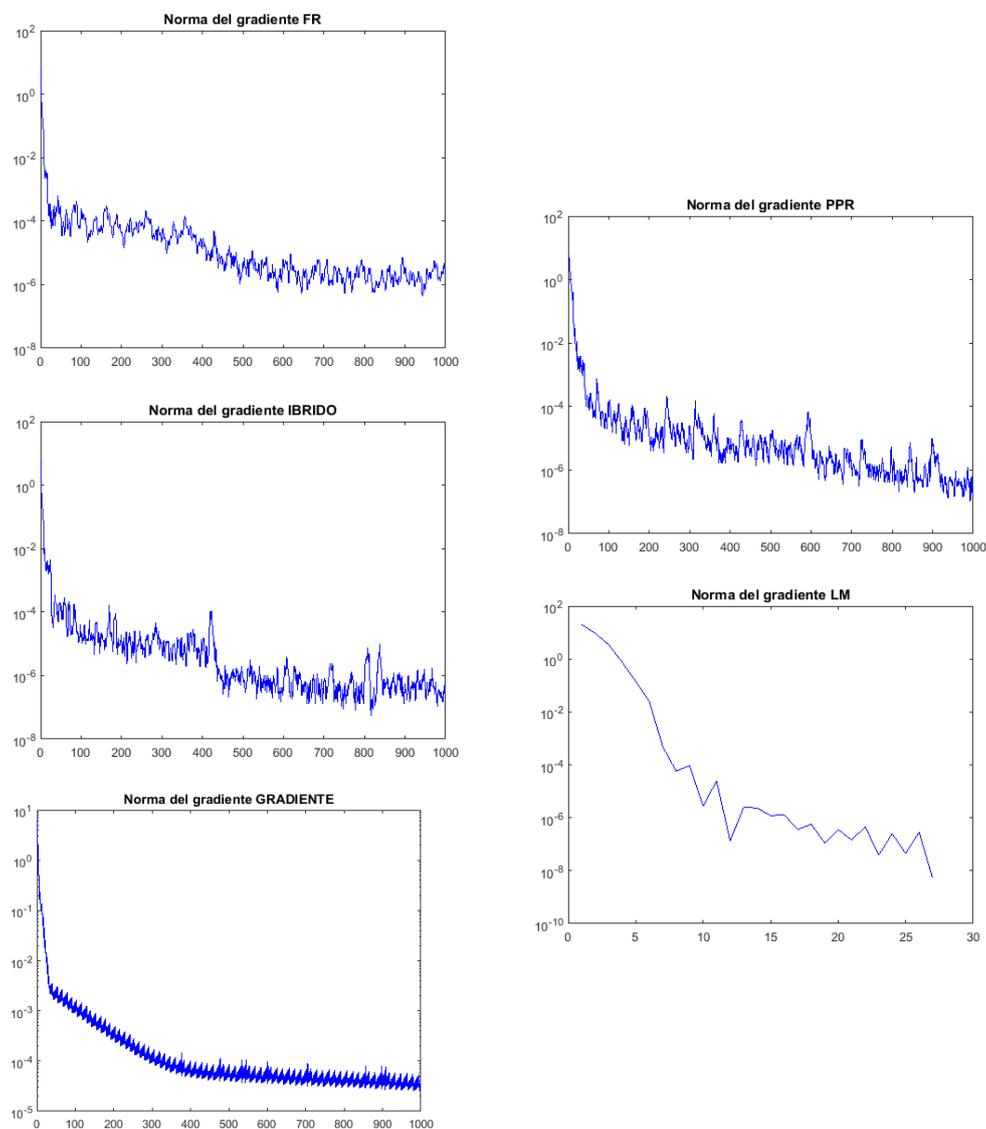


Figura 5.1: norma del gradiente, test 1, senza rumore

Per quanto riguarda le tre versioni non lineari del CG, i risultati confermano quanto visto nella teoria: il metodo FR come metodo di ottimizzazione è il più lento, infatti dopo 1000 iterazioni la norma del gradiente è ancora dell'ordine di 10^{-6} ; a differenza, invece, del metodo PPR e del metodo IBRIDO, con i quali si raggiunge una norma del gradiente dell'ordine di 10^{-7} . In ogni

caso, i metodi risultano essere più veloci del metodo del gradiente standard, mentre, come ci si aspettava, il metodo di Lavenberg-Marquardt, essendo un metodo di ottimizzazione di secondo grado, è molto più veloce.

Consideriamo adesso il caso con rumore.

`noise=0.01, 'seed'=0,`

Nella tabella it_{min} rappresenta l'iterazione corrispondente al minimo errore relativo e it_{dis} l'iterazione alla quale viene arrestato il metodo utilizzando il principio della discrepanza con $\tau = 1$.

	Minimo errore relativo	it_{min}	Errore relativo (discrepanza)	it_{dis}
CG FR	7.93e-03	8	7.93e-03	8
CG PPR	7.82e-03	43	9.59e-03	15
CG IBRIDO	7.93e-03	8	7.93e-03	8
Lavenberg-Marquardt	9.02e-03	8	1.03e-02	7
Gradiente	8.78e-03	236	1.50e-02	16

Si nota che il metodo di Lavenberg-Marquardt, pur essendo il metodo di ottimizzazione più veloce, produce un errore relativo minimo più grande rispetto agli altri metodi. Il metodo PPR produce il più piccolo errore relativo minimo, ma dopo ben 43 iterazioni, a differenza del metodo FR e del metodo IBRIDO che, come il metodo di Lavenberg-Marquardt, impiegano solo 8 iterazioni per raggiungere l'errore relativo minimo. Per quanto riguarda il metodo del gradiente, esso è il più lento nella semiconvergenza: impiega 236 iterazioni per raggiungere l'errore relativo minimo, che in ogni caso risulta essere superiore a quello raggiunto dai CG.

È possibile inoltre osservare dalla tabella, che utilizzando come criterio d'arresto il principio della discrepanza, il metodo FR e il metodo IBRIDO producono esattamente la soluzione corrispondente all'errore relativo minimo. Per quanto riguarda il metodo PPR viene arrestato ben 28 iterazioni prima con un errore relativo che si discosta dall'errore relativo minimo di

circa $1.5e-03$, a differenza invece del metodo di Lavenberg-Marquardt il quale, con il principio della discrepanza, viene arrestato all'iterazione precedente producendo però un errore relativo di $1e-03$ più grande rispetto al minimo.

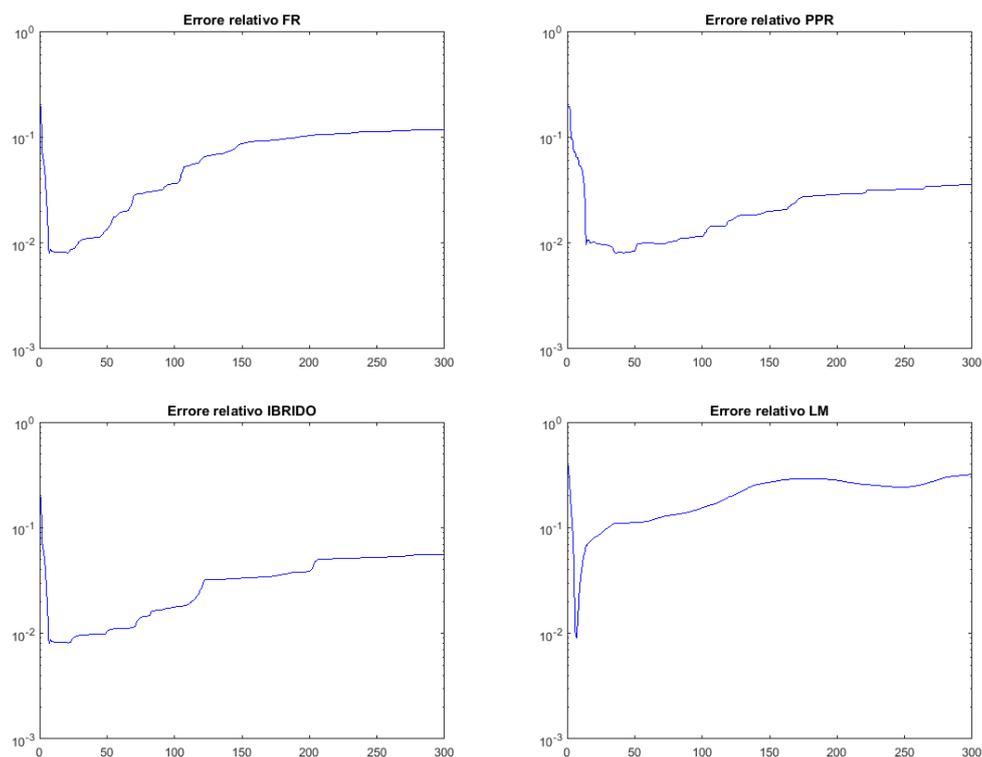


Figura 5.2: Semiconvergenza dei metodi

Osservando i grafici dell'errore relativo è evidente, innanzitutto, la semiconvergenza delle versioni del CG, i quali per tanto possono essere identificati nella categoria dei metodi di regolarizzazione iterativi. Inoltre i loro bacini di semiconvergenza sono notevolmente più piatti rispetto a quello di Lavenberg Marquardt: il metodo FR e l'IBRIDO raggiungono l'errore minimo relativo con lo stesso numero di iterazioni di Lavenberg-Marquardt, ma semiconvergono molto più lentamente. Questo è un'aspetto vantaggioso; la scelta dell'iterazione alla quale arrestare metodi di questo tipo è di più facile gestione: anche se il metodo viene fermato 10 iterazioni più avanti l'errore relativo non si allontanerà di molto da quello minimo.

noise=0.1, 'seed'=0,

	Minimo errore relativo	it_{min}	Errore relativo (discrepanza)	it_{dis}
CG FR	3.70e-02	11	9.05e-02	3
CG PPR	3.74e-02	18	8.56e-02	8
CG IBRIDO	3.64e-02	9	9.05e-02	3
Lavenberg-Marquardt	4.64e-02	7	1.06e-01	5
Gradiente	3.93e-02	14	8.56e-02	5

Aumentando il rumore aumenta la differenza tra l'errore relativo minimo calcolato da Lavenberg-Marquardt e quello invece calcolato dalle varie versioni del CG; mentre con un rumore dell'ordine di 10^{-2} i due errori differivano tra loro di un fattore dell'ordine di 10^{-3} , ora differiscono di circa $2 \cdot 10^{-2}$. In questo caso il metodo di tipo CG che produce un errore relativo minimo migliore è il metodo IBRIDO.

Anche aumentando il grado del rumore, l'arresto con il principio della discrepanza produce risultati migliori per i metodi di tipo CG. Questo comportamento si riconduce anche in questo caso alla velocità di semiconvergenza dei metodi. Il metodo di Lavenberg-Marquardt, anche se viene arrestato solo 2 iterazioni prima produce un errore relativo di circa $6.6e-02$ più grande; ciò è indice del fatto che semiconverge troppo velocemente, a differenza degli altri metodi i quali, una volta raggiunto l'errore minimo, si allontanano dalla soluzione esatta abbastanza lentamente.

Visualizziamo gli iterati calcolati relativi all'errore relativo minimo a confronto con la soluzione esatta (linea rossa), rispettivamente del metodo IBRIDO che produce il più piccolo errore relativo minimo e il metodo di Lavenberg Marquardt che invece produce il più grande errore relativo minimo.

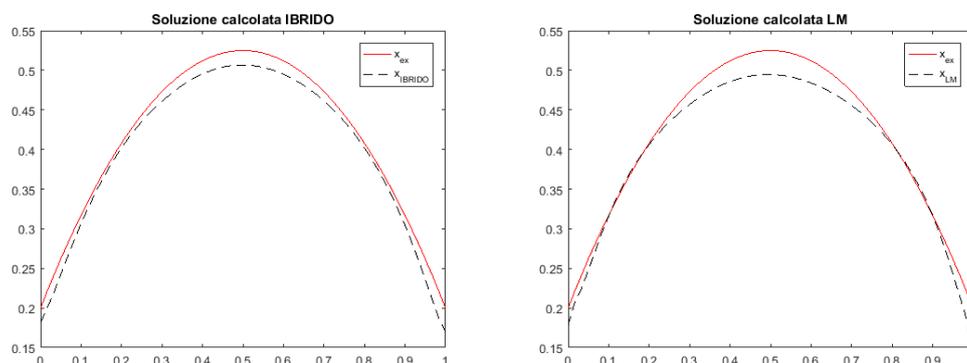


Figura 5.3: Soluzione test 1, rumore 0.1 , 60 punti di discretizzazione

Consideriamo adesso, ponendo sempre $\text{noise}=0.1$, discretizzazioni differenti dell'intervallo. I risultati che si ottengono sono i seguenti:

$$m = n = 10$$

	Minimo errore relativo	it_{min}	norma gradiente (dopo 1000 iterazioni)	Errore relativo (discrepanza)	it_{dis}
FR	5.19e-02	9	1.24e-02	9.72e-02	3
PPR	5.82e-02	23	1.61e-03	9.81e-02	8
IBRIDO	5.71e-02	7	1.50e-05	9.72e-02	3
LM	7.25e-02	5	2.79e-11	1.03e-01	4
GRAD	6.31e-02	28	2.93e-03	9.81e-02	5

$$m = n = 120$$

	Minimo errore relativo	it_{min}	norma gradiente (dopo 1000 iterazioni)	Errore relativo (discrepanza)	it_{dis}
FR	3.41e-02	9	1.00e-01	8.96e-02	3
PPR	3.24e-02	19	9.82e-04	8.49e-02	8
IBRIDO	3.41e-02	9	2.32e-03	8.96e-02	3
LM	3.82e-02	8	8.62e-04	1.024e-01	6
GRAD	3.55e-02	30	2.13e-03	8.49e-02	5

Osservando la norma del gradiente dopo mille iterazioni si può osservare che con una discretizzazione meno fine i metodi sono più veloci, tuttavia la soluzione calcolata risulta essere più sensibile all'aumento del rumore, in quanto l'errore relativo minimo è più grande rispetto al caso di $m = n = 120$.

Per quanto riguarda l'arresto con il principio della discrepanza si conferma quanto detto precedentemente.

Visualizziamo anche in questo caso il grafico della soluzione calcolata dal metodo di tipo CG che produce il minore errore relativo minimo e dal metodo di Lavenberg-Marquardt, confrontate con la soluzione esatta (linea rossa) rispettivamente con discretizzazioni degli intervalli scelte.

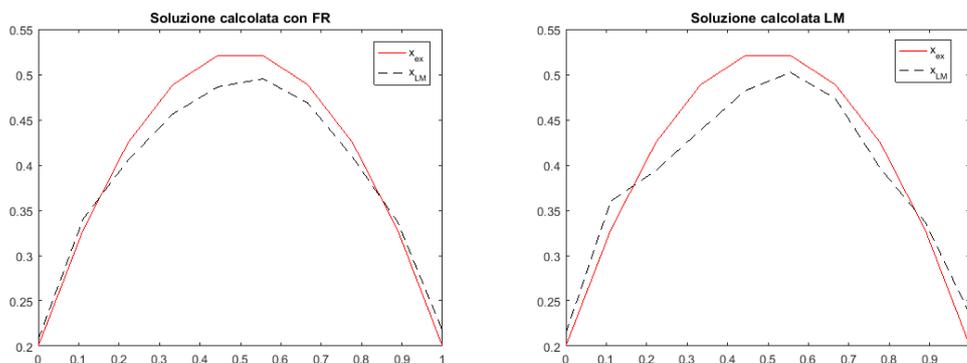


Figura 5.4: Soluzione test 1, rumore 0.1 , 10 punti di discretizzazione

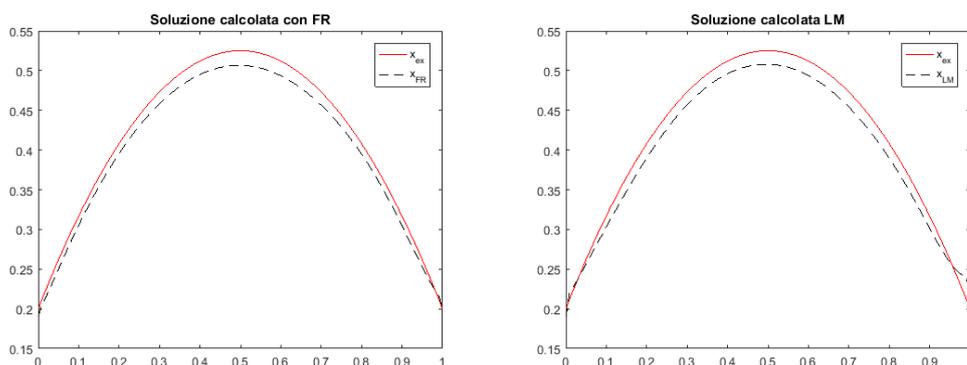


Figura 5.5: Soluzione test 1, rumore 0.1 , 120 punti di discretizzazione

- $x_0 = 0.5 * \text{ones}(\text{size}(x_{ex}))$

$m = n = 60$

$\text{noise}=0.01$, 'seed'=0,

	Minimo errore relativo	it_{min}	Errore relativo (discrepanza)	it_{dis}
CG FR	8.52e-02	51	1.13e-01	6
CG PPR	8.54e-02	32	1.15e-01	14
CG IBRIDO	8.67e-02	21	1.13e-01	6
Lavenberg-Marquardt	8.84e-02	7	1.58e-01	4
Gradiente	9.19e-02	141	1.59e-01	10

Scegliendo questo iterato l'errore relativo minimo calcolato da tutti i metodi è superiore rispetto al caso precedente, e i metodi di tipo CG hanno bisogno di più iterazioni per raggiungere l'errore minimo, che in ogni caso risulta essere inferiore a quello di Lavenberg-Marquardt.

A questo punto sempre con lo stesso iterato iniziale abbiamo provato a variare il 'seed' e per alcuni valori di esso si è visto che i metodi di tipo CG possono semiconvergere più lentamente; ne riportiamo un esempio:

$\text{noise}=0.01$, 'seed'=-2,

	Minimo errore relativo	it_{min}	Errore relativo (discrepanza)	it_{dis}
CG FR	6.41e-02	404	8.88e-02	17
CG PPR	7.13e-02	1001	9.76e-02	30
CG IBRIDO	7.23e-02	808	9.71e-02	13
Lavenberg-Marquardt	7.39e-02	17	9.11e-02	6
Gradiente	9.17e-02	1001	9.84e-02	30

Come si può vedere dalla tabella tutti i metodi hanno bisogno di un gran numero di iterazioni per raggiungere l'errore relativo minimo, addirittura il metodo PPR, così come il metodo del gradiente, in 1000 iterazioni non semiconverge. Tuttavia arrestando i metodi con il principio della discrepanza

vengono arrestati dopo massimo 30 iterazioni e anche questa volta il metodo FR produce la soluzione più vicina. Visualizziamo le soluzioni calcolate dopo l'arresto con il principio della discrepanza.

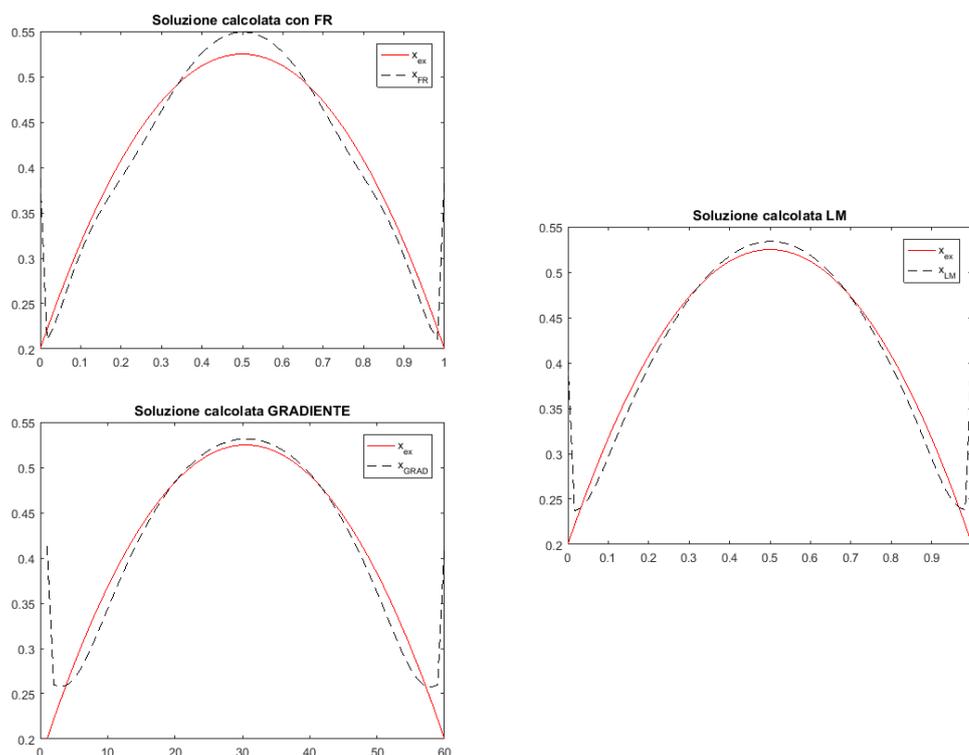


Figura 5.6: Soluzione test 1, rumore 0.1 , seed=-2

Questo comportamento invece non si ha scegliendo 'seed'=-2 e noise=0.1

	Minimo errore relativo	it_{min}	Errore relativo (discrepanza)	it_{dis}
CG FR	9.78e-02	19	1.61e-01	5
CG PPR	1.01e-01	41	1.15e-01	13
CG IBRIDO	9.94e-02	21	1.62e-01	5
Lavenberg-Marquardt	1.032e-01	6	1.55e-01	4
Gradiente	1.05e-01	77	1.53e-01	9

In questo caso seppure il metodo PPR non raggiunga l'errore minimo più basso, arrestandolo con il principio della discrepanza produce la soluzione

più vicina a quella esatta.

- $x_0 = \text{ones}(\text{size}(\mathbf{x}_{\text{ex}}))$

noise=0.01, 'seed'=0

	Minimo errore relativo	it_{min}	norma gradiente (dopo 1000 iterazioni)	Errore relativo (discrepanza)	it_{dis}
FR	1.29+00	3	1.07e-03	1.56+00	12
PPR	8.56e-02	43	1.89e-04	1.63e-01	43
IBRIDO	1.30+00	3	8.82e-05	1.56e+00	10
LM	1.35e-01	99	9.95e-07	2.37e-01	12
GRAD	6.76e-02	401	5.35e-04	8.37e-02	83

Con questa scelta dell'iterato iniziale notiamo che i metodi FR e IBRIDO non convergono alla soluzione esatta; visualizzando, infatti, le loro rispettive soluzioni calcolate all'iterazione corrispondente all'errore relativo minimo si osserva che questa è molto più vicina all'altra soluzione esatta del problema, che ricordiamo essere la funzione: $x(s) = 1.3s(s - 1)$:

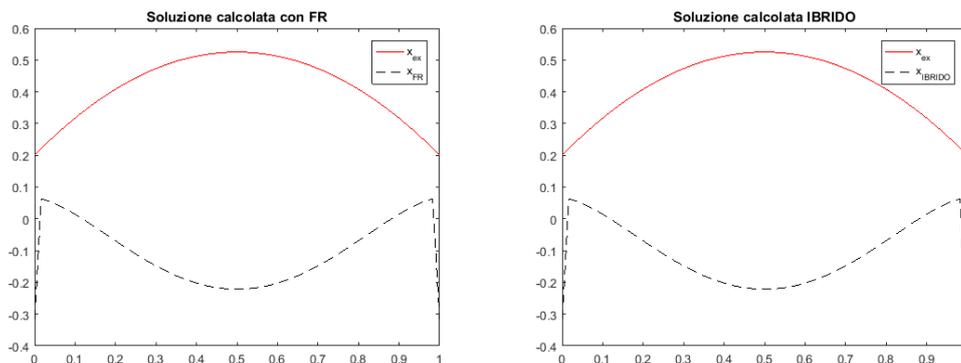


Figura 5.7: Soluzione test 1, FR e IBRIDO , $x_0 = \text{ones}(\text{size}(\mathbf{x}_{\text{ex}}))$

Il metodo PPR si comporta meglio del metodo di Lavenberg-Marquardt, in quanto ha bisogno di un minor numero di iterazioni per semiconvergere e raggiunge un errore minimo relativo più basso. Anche in questo caso il

metodo del gradiente viene arrestato dopo un numero di iterazioni maggiore rispetto agli altri metodi, ma in termini di errore relativo la sua performance è la migliore.

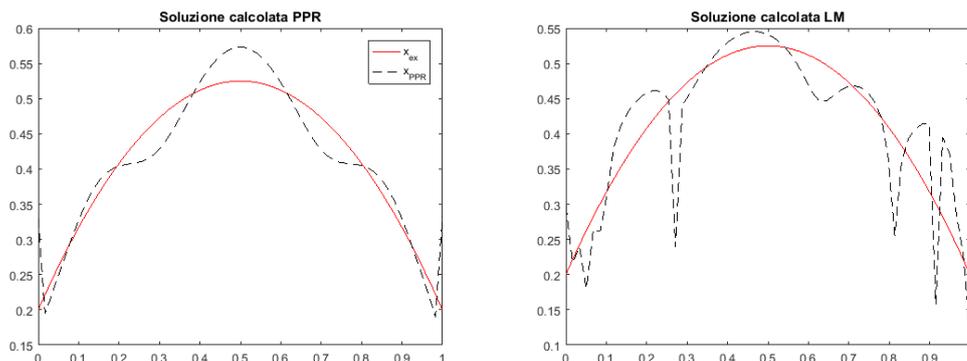


Figura 5.8: Soluzione test 1, PPR e LM, $x_0 = \text{ones}(\text{size}(x_{ex}))$

5.2 Problema test 2

Come secondo problema test consideriamo sempre l'equazione di Fredholm (5.1) con kernel (5.2) e $H=0.2$. Tale problema, analizzato in [10] ammette due soluzioni continue $x_{ex}(s) = 2H - c_1 e^{d_1(s+p_1)^2} - c_2 e^{d_2(s-p_2)^2} - c_3 s - c_4$ e $x_{ex}(s) = c_1 e^{d_1(s+p_1)^2} + c_2 e^{d_2(s-p_2)^2} + c_3 s + c_4$ con $c_1 = -0.1$, $c_2 = -0.075$, $d_1 = -40$, $d_2 = -60$, $p_1 = 0.4$, $p_2 = 0.67$ e c_3 e c_4 scelti in modo che $x_{ex}(0) = x_{ex}(1) = 0$.

Il problema è stato implementato allo stesso modo del problema test 1, e come soluzione esatta si è scelta la funzione $x_{ex}(s) = c_1 e^{d_1(s+p_1)^2} + c_2 e^{d_2(s-p_2)^2} + c_3 s + c_4$. Anche in questo caso sono stati fatti diversi test variando l'iterato iniziale x_0 , la discretizzazione degli intervalli e il rumore. Riportiamo qui di seguito alcuni dei risultati più significativi.

- $x_0 = \text{zeros}(\text{size}_{x_{ex}})$

$$m = n = 60, \text{noise}=0$$

	Errore relativo	norma del gradiente	iter
CG FR	4.13e-02	1.16e-05	1000
CG PPR	6.00e-02	2.23e-05	1000
CG IBRIDO	4.63e-02	9.64e-06	1000
Lavenberg-Marquardt	7.87e-03	6.15e-09	19
Gradiente	1.19e-01	2.71e-04	1000

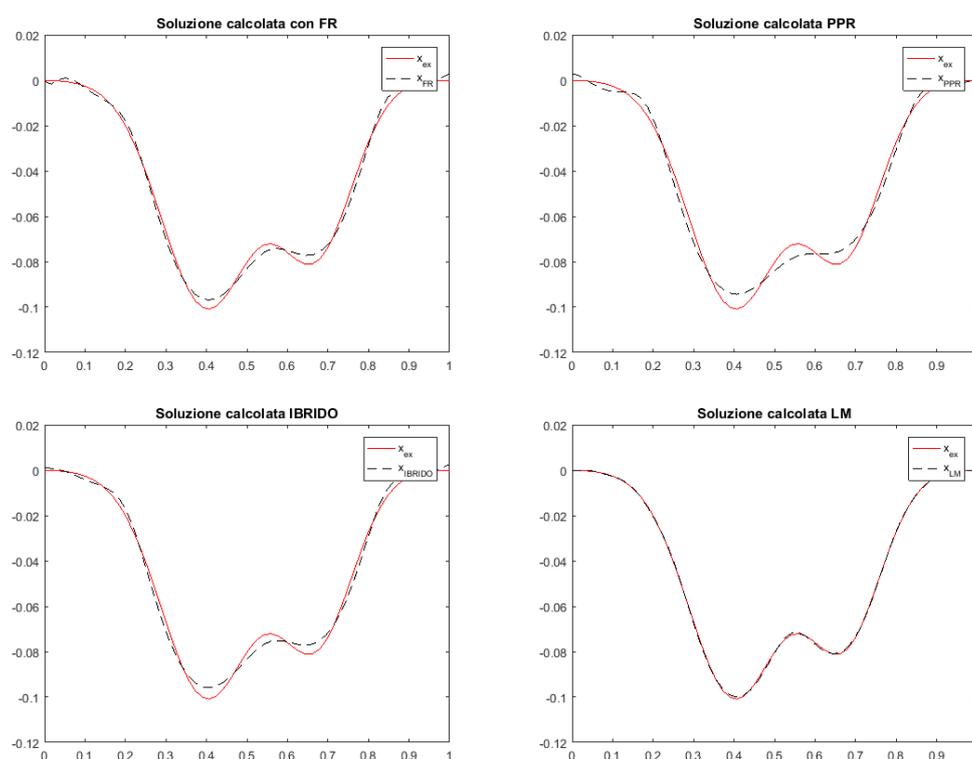
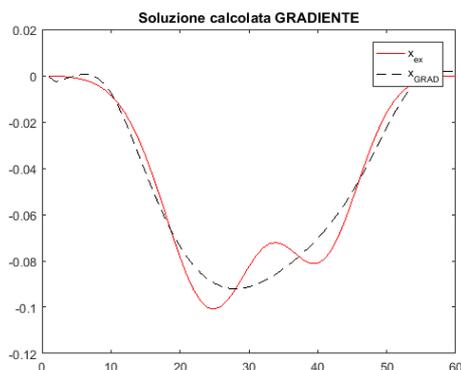


Figura 5.9: Soluzione test 2

Anche in questo esempio è possibile notare come il metodo di Lavenberg-Marquardt risulta essere il metodo di ottimizzazione più veloce, a seguire i tre metodi di tipo CG e infine il gradiente che risulta essere il più lento. Come già precedentemente discusso, però, nei problemi pratici bisogna mettere in conto una quantità, seppur minima, di rumore; ripetiamo dunque lo stesso esperimento aggiungendo del rumore e arrestando il metodo con il principio della discrepanza; questa volta si è scelto $\tau = 0.5$.



Come prima ci aspettiamo che i metodi semiconvercano, analizziamo in che modo:

`noise=0.01, 'seed'=0`

	Minimo errore relativo	it_{min}	Errore relativo (discrepanza)	it_{dis}
CG FR	1.65e-01	15	1.65e-01	13
CG PPR	1.60e-01	33	1.61e-01	12
CG IBRIDO	1.65e-01	16	1.65e-01	14
Lavenberg-Marquardt	1.57e-01	5	1.57e-01	5
Gradiente	1.62e-01	104	1.63e-01	9

Dalla tabella si nota che, innanzitutto gli errore relativi minimi raggiunti dai metodi sono tutti più alti rispetto al problema test precedente, seppure stiamo considerando la stessa quantità di rumore; questo vuol dire che il livello di mal condizionamento del problema è superiore rispetto al problema precedente. L'errore relativo minimo più piccolo questa volta viene raggiunto dal metodo di Lavenberg Marquardt; a seguire il metodo PPR.

Visualizzando i grafici degli errori si nota, anche in questo caso, come i metodi di tipo CG abbiano un bacino di semiconvergenza più piatto rispetto al metodo di Lavenberg Marquardt, e di come dunque, sia più facile scegliere l'iterazione opportuna dove fermare il metodo.

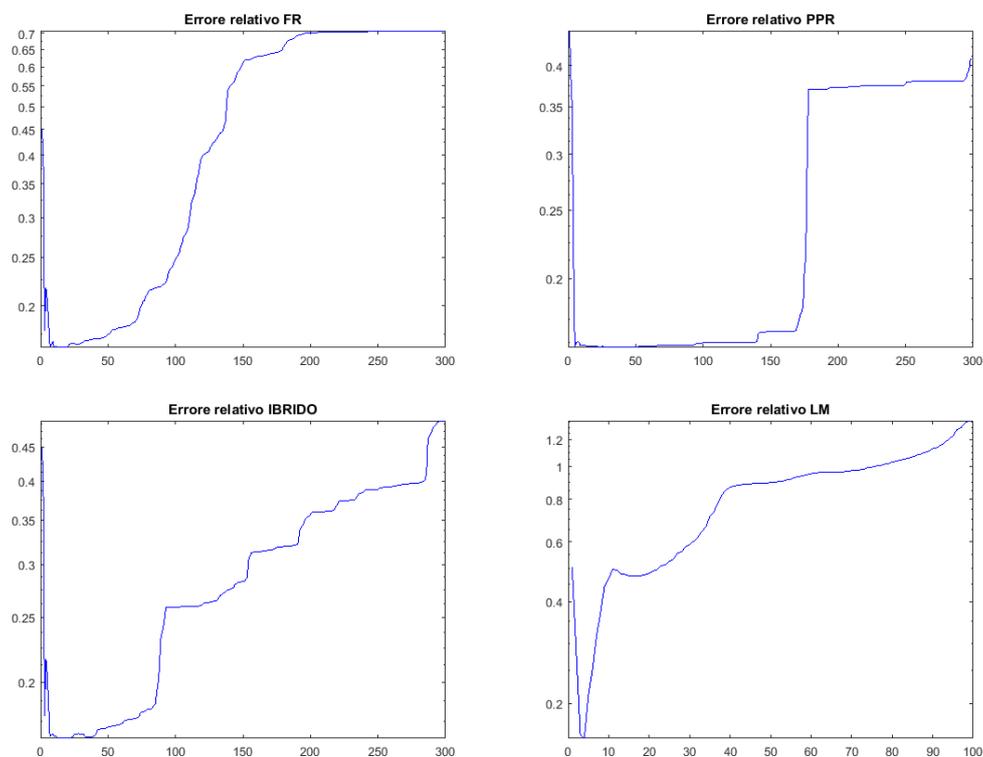


Figura 5.10: Semiconvergenza test 2, rumore 0.01

noise=0.1, 'seed'=0

	Minimo errore relativo	it_{min}	Errore relativo (discrepanza)	it_{dis}
CG FR	2.34e-01	4	3.10e-01	7
CG PPR	2.02e-01	3	2.80e-01	4
CG IBRIDO	2.34e-01	4	2.89e-01	8
Lavenberg-Marquardt	3.32e-01	3	3.61e-01	4
Gradiente	2.90e-01	8	3.00e-01	7

In questo caso il metodo che si comporta meglio è il metodo PPR che raggiunge nello stesso numero di iterazioni di Lavenberg-Marquardt l'errore relativo minimo più basso.

Le figure che seguono rappresentano, rispettivamente l'errore relativo del

metodo PPR (in rosso) e quello del metodo di Lavenberg Marquardt (in blu) alle prime iterazioni e le soluzioni migliori calcolate dai due metodi:

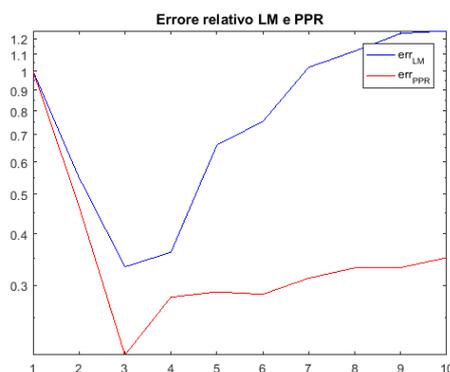


Figura 5.11: Semiconvergenza test 2, PPR e LM, rumore 0.1

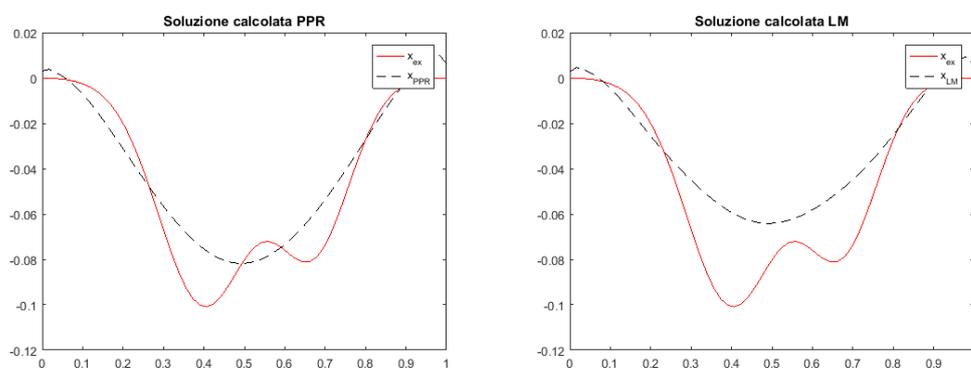


Figura 5.12: Soluzioni test 2, PPR e LM, rumore 0.1

Come si può notare le soluzioni calcolate, soprattutto quella di Lavenberg-Marquardt non sono molto soddisfacenti. Anche in questo caso si è provato ad aumentare i punti di discretizzazione scegliendo $m = n = 120$, ottenendo risultati leggermente migliori:

	Minimo errore relativo	it_{min}	Errore relativo (discrepanza)	it_{dis}
CG FR	2.26e-01	4	3.04e-01	7
CG PPR	1.97e-01	3	2.73e-01	4
CG IBRIDO	2.34e-01	4	2.80e-01	8
Lavenberg-Marquardt	2.76e-01	4	2.76e-01	4
Gradiente	2.74e-01	8	2.84e-01	7

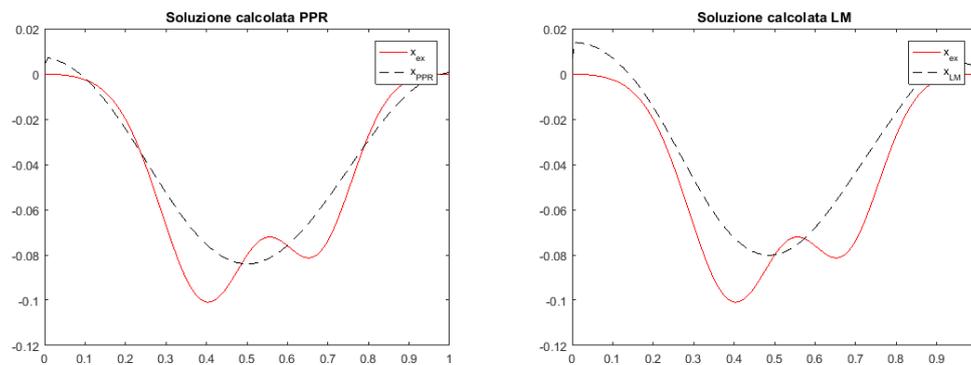


Figura 5.13: Soluzioni test 2, PPR e LM, rumore 0.1, 120 punti

Proviamo a questo punto a scegliere un iterato iniziale ancora più lontano dalla soluzione esatta:

- $x_0 = 0.1 * \text{ones}(\text{size}(x_{ex}))$

$m = n = 60$, $\text{noise}=0$

	Errore relativo	norma del gradiente	iter
CG FR	1.24e-01	2.90e-05	1000
CG PPR	1.25e-01	9.00e-06	1000
CG IBRIDO	1.21e-01	5.71e-06	1000
Lavenberg-Marquardt	8.20e-02	6.61e-09	54
Gradiente	1.72e-01	3.41e-04	1000

Scegliendo iterati più distanti dalla soluzione esatta accade che, in termini di norma del gradiente della funzione obiettivo, i metodi hanno più o meno

gli stessi risultati, ma la soluzione calcolata si allontana maggiormente dalla soluzione esatta.

Visualizzando le soluzioni ottenute, ad esempio quella del metodo IBRIDO e quella di Lavenberg-Marquardt si nota che le soluzioni calcolate dai metodi si allontanano maggiormente dalla soluzione esatta nei pressi degli estremi dell'intervallo $[0, 1]$.

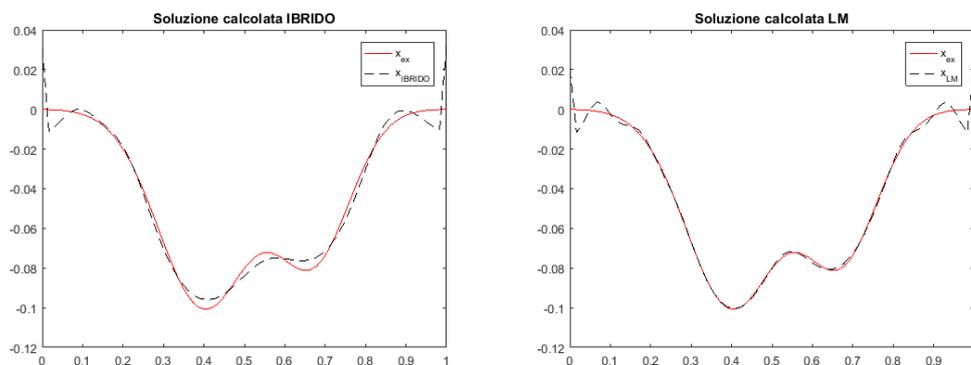


Figura 5.14: Soluzioni test 2, IBRIDO e LM, $\mathbf{x}_0 = 0.1 * \text{ones}(\text{size}(\mathbf{x}_{\text{ex}}))$

`noise=0.01, 'seed'=0`

	Minimo errore relativo	it_{min}	Errore relativo (discrepanza)	it_{dis}
CG FR	1.95e-01	128	2.30e-01	15
CG PPR	1.99e-01	177	2.44e-01	23
CG IBRIDO	2.03e-01	83	2.32e-01	16
Lavenberg-Marquardt	2.09e-01	7	2.09e-01	7
Gradiente	2.11e-01	1001	2.53e-01	102

Dunque con questo particolare iterato iniziale i metodi di tipo CG semi-convergono tutti molto più lentamente del metodo di Lavenberg Marquardt.

La seguente figura mostra il grafico dell'errore relativo del metodo IBRIDO (in rosso) e del metodo di Lavenberg-Marquardt (in blu) alle prime 100 iterazioni:

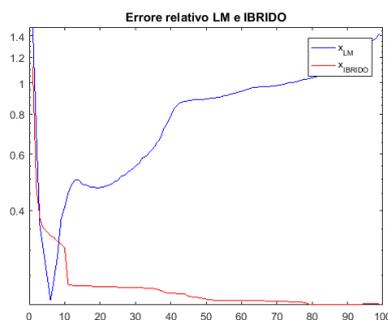


Figura 5.15: Errore relativo test 2, LM e IBRIDO, rumore 0.01

Come si può osservare dal grafico, il metodo IBRIDO raggiunge l'errore minimo relativo più lentamente del metodo di Lavenberg-Marquardt che invece impiega solo 7 iterazioni. Tuttavia, a causa della rapidità della sua semiconvergenza, arrestandolo all'iterazione successiva si ottiene una soluzione con un errore relativo pari alla soluzione ottenuta arrestando il metodo IBRIDO all'iterazione $k=12$ ed inoltre la qualità delle due soluzioni sarà differente: mentre nel caso di Lavenberg-Marquardt l'arresto viene effettuato ad una iterazione in cui il metodo di sta allontanando dalla soluzione esatta nel caso dell'IBRIDO si sta avvicinando.

Visualizziamo rispettivamente la soluzione calcolata dal metodo IBRIDO all'iterazione $k=12$ e la soluzione calcolata da Lavenberg-Marquardt all'iterazione $k=8$.

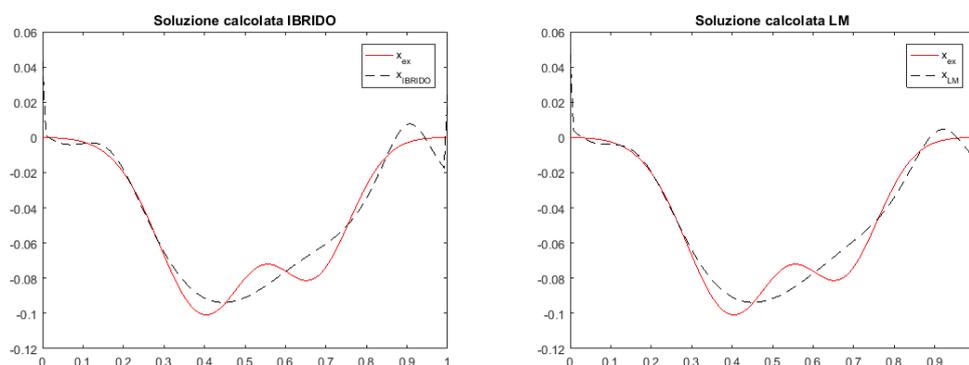


Figura 5.16: Soluzioni test 2, IBRIDO e LM, $x_0 = 0.1 * \text{ones}(\text{size}(x_{ex}))$

noise=0.1, 'seed'=0

	Minimo errore relativo	it_{min}	Errore relativo (discrepanza)	it_{dis}
CG FR	3.69e-01	4	4.47e-01	5
CG PPR	3.53e-01	4	4.09e-01	5
CG IBRIDO	3.69e-01	4	4.46e-01	5
Lavenberg-Marquardt	4.12e-01	5	4.12e-01	5
Gradiente	3.53e-01	3	4.09e-01	4

Visualizziamo il grafico dell'errore relativo per analizzare la semiconvergenza.

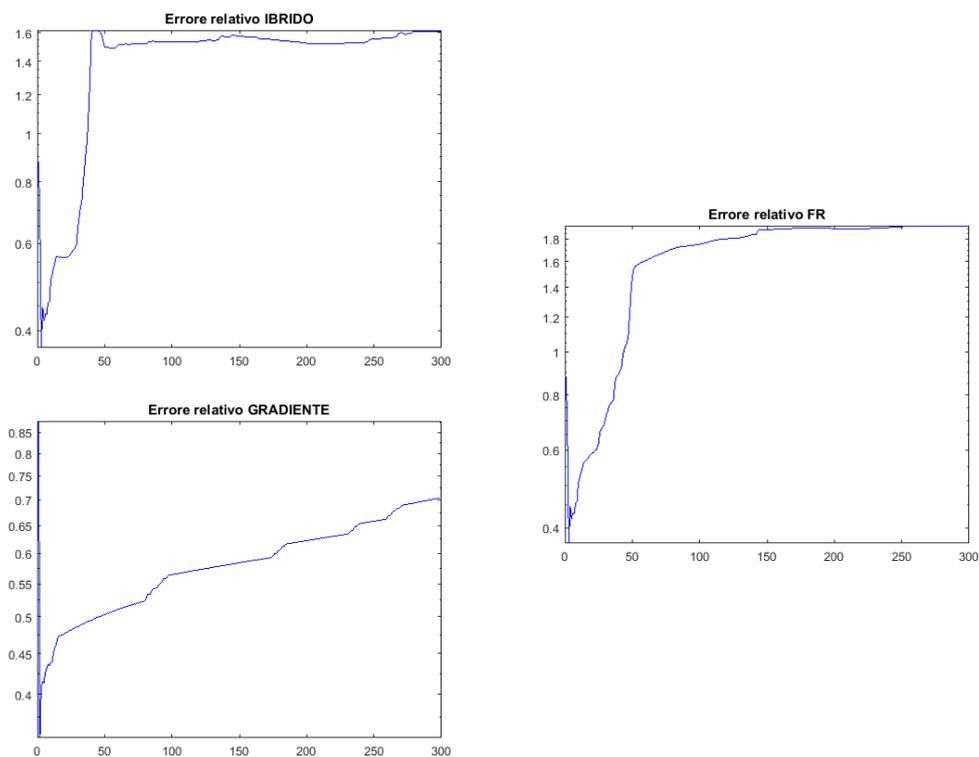


Figura 5.17: Semiconvergenza test 2, rumore 0.1

Aumentando il grado del rumore aumenta anche la velocità di semiconvergenza dei metodi. Il metodo di Lavenberg-Marquardt rimane ancora il

metodo che semiconverge più velocemente. Nella figura accanto viene confrontato il grafico del suo errore relativo (blu) con quello del metodo PPR (rosso).

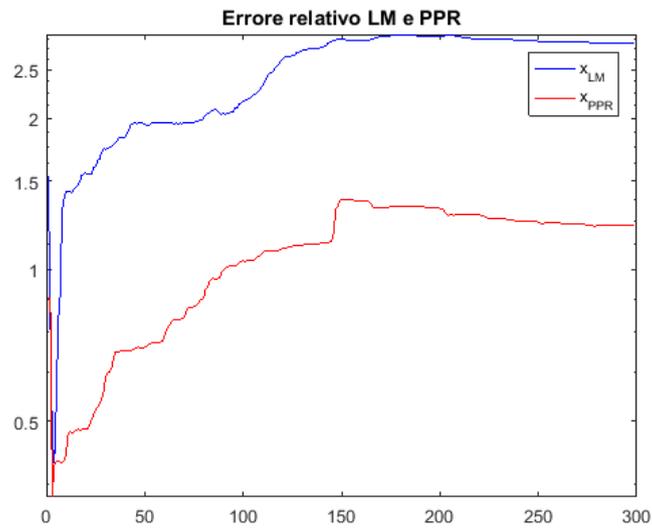


Figura 5.18: Semiconvergenza test 2, rumore 0.1, LM e PPR

Conclusioni

In questo elaborato è stato analizzato il metodo del gradiente coniugato, nato appositamente per la risoluzione di sistemi lineari o equivalentemente per l'ottimizzazione di funzioni quadratiche convesse e sono state considerate tre particolari versioni del metodo per il caso non quadratico, analizzandone la loro applicazione a problemi ai minimi quadrati non lineari mal condizionati. Come ci si aspettava dai risultati teorici, tali metodi, se utilizzati per l'ottimizzazione di funzioni, risultano essere più lenti rispetto ai metodi di tipo Newton, anche se più veloci rispetto al metodo del gradiente. Tuttavia, come è emerso dai due problemi test, questi metodi hanno fornito buoni risultati nell'ambito della regolarizzazione di problemi ai minimi quadrati non lineari mal condizionati e sono un'interessante alternativa ai metodi di regolarizzazione iterativi presenti in letteratura (Gauss-Newton, Lavenberg-Marquardt). In particolare sono state messe in evidenza le buone proprietà di regolarizzazione dei metodi: essi raggiungono un errore relativo minimo più basso sia rispetto al metodo di Lavenberg-Marquardt, sia rispetto al metodo del gradiente, in un numero di iterazioni confrontabile con il metodo di Lavenberg-Marquardt. Il metodo del gradiente si avvicina alla soluzione esatta lentamente e altrettanto lentamente se ne allontana al crescere delle iterazioni; le tre versioni del gradiente coniugato non lineare, invece, si avvicinano alla soluzione esatta più velocemente rispetto al metodo del gradiente e se ne allontanano più lentamente rispetto al metodo di Lavenberg-Marquardt. È stato inoltre analizzato l'utilizzo del principio della discrepanza come criterio di arresto per i metodi presi in esame e anche in questo caso sono stati

ottenuti risultati soddisfacenti: si è osservato infatti che, arrestando tutti i metodi con il principio della discrepanza, la soluzione ottenuta con la versione del metodo del gradiente non lineare è di qualità maggiore rispetto a quelle ottenute con il metodo del gradiente e di Lavenberg-Marquardt. Complessivamente tutti e tre i metodi esaminati si sono rivelati essere buoni metodi di regolarizzazione e non sono state riscontrate differenze evidenti tra loro: in alcuni casi la soluzione più vicina è stata calcolata dal metodo PPR, in altri dal metodo FR, entrambi con più o meno lo stesso numero di iterazioni. Tuttavia possiamo considerare il metodo ibrido come il più robusto tra i tre.

Dati i risultati positivi ottenuti in questo lavoro, sarebbe interessante proseguire nello studio di questi metodi e analizzarne l'applicazione a problemi ai minimi quadrati non lineari più complessi, nell'ambito, ad esempio, di tecniche di tomografia avanzate.

Bibliografia

- [1] G. Scherer, V. Pereya, C. Hansen, *Least squares data fitting with applications* (2012)
- [2] S. Bellavia, B. Morini, E. Riccietti *On an adaptive regularization for ill-posed non linear systems and its trust-region implementation* (2015)
- [3] L. Grippo, M. Sciandrone, *Metodi di ottimizzazione non vincolata*, Springer (2011)
- [4] Y. Yuan, *Analysis on the conjugate gradient method*, Optim. Methods Softw., 2 (1993) pp 19-29
- [5] J. D. Powell, *Nonconvex minimization calculations and the conjugate gradient method*, Numerical Analysis (Dundee,1983), Lecture Notes in Mathematics, Vol.1066, Springer-Verlag, Berlin, (1984), pp 122-141.
- [6] Wang, Y., Yuan, Y., *Convergence and regularity of trust region methods for nonlinear ill-posed inverse problems* Inverse Prob. 21, pp 821-838 (2005)
- [7] M. J. D. Powell, *Restart procedures of the conjugate gradient method*, Math. Prog., 2 (1977), pp. 241-254.
- [8] D. Touati-Ahmed and C.Storey, *Efficient hybrid conjugate gradient techniques*, J. Optim. Theory Appl., 64 (1990), pp. 379-397
- [9] J. C. Gilbert and J. Nocedal, *Global convergence properties of conjugate gradient methods for optimization*, SIAM J. Optim., 2 (1992)

- [10] Vogel, C. R. *A constrained least squares regularization method for nonlinear ill-posed problems*. SIAM J. Control Optim. 28, 34-49 (1990)
- [11] Per Christian Hansen, *Rank-Deficient and Discrete Ill-Posed Problems*, SIAM, (1998)
- [12] J. Nocedal, S. J. Wright, *Numerical Optimization*, Springer, (2006)
- [13] V. Comincioli, *Analisi numerica* McGraw-Hill, Milano, (1990)