

Luigi Grippo · Marco Sciandrone

Metodi di ottimizzazione non vincolata



Springer

Indice

Prefazione	XIII
1 Problemi di ottimizzazione su R^n	1
1.1 Generalità	1
1.2 Definizioni fondamentali	5
1.3 Criteri elementari di equivalenza tra problemi	10
1.4 Condizioni di esistenza	12
1.5 Formulazione dei problemi di ottimo non vincolati	20
1.5.1 Equazioni e disequazioni	21
1.5.2 Stima dei parametri di un modello matematico	22
1.5.3 Addestramento di reti neurali	23
1.5.4 Problemi di controllo ottimo	26
1.5.5 Funzioni di penalità sequenziali	27
1.5.6 Proprietà delle funzioni di penalità sequenziali*	28
1.6 Esercizi	32
2 Condizioni di ottimo per problemi non vincolati	35
2.1 Generalità	35
2.2 Direzioni di discesa	36
2.3 Condizioni di ottimalità	39
2.3.1 Condizioni di minimo locale	39
2.3.2 Condizioni di minimo globale nel caso convesso	44
2.3.3 Condizioni di ottimo in problemi di minimi quadrati ...	46
2.4 Equazioni non lineari	49
2.5 Esercizi	51
3 Struttura e convergenza degli algoritmi	55
3.1 Generalità	55
3.2 Punti di accumulazione	57
3.3 Convergenza a punti stazionari	64
3.4 Rapidità di convergenza	69

3.5	Classificazione degli algoritmi convergenti	71
3.6	Esercizi	74
4	Convergenza di metodi con ricerche unidimensionali	75
4.1	Generalità	75
4.2	Condizioni di convergenza globale: metodi monotoni	76
4.3	Condizioni di convergenza globale: metodi non monotoni*	84
4.4	Esercizi	86
5	Ricerca unidimensionale	87
5.1	Generalità	87
5.1.1	Ricerca di linea esatta	89
5.1.2	Ricerche di linea inesatte	90
5.2	Metodo di Armijo	92
5.2.1	Definizione del metodo e convergenza	92
5.2.2	Estensioni dei risultati di convergenza*	100
5.2.3	Metodo di Armijo con gradiente Lipschitz-continuo*	102
5.3	Tecniche di espansione, condizioni di Goldstein	106
5.4	Metodo di Wolfe	111
5.4.1	Condizioni di Wolfe e convergenza	111
5.4.2	Metodo di Wolfe con gradiente Lipschitz-continuo	115
5.4.3	Algoritmi basati sulle condizioni di Wolfe*	116
5.5	Ricerca unidimensionale senza derivate	123
5.6	Ricerca unidimensionale non monotona	132
5.6.1	Metodo di Armijo non monotono	132
5.6.2	Ricerca unidimensionale non monotona: convergenza*	135
5.7	Realizzazione di algoritmi di ricerca unidimensionale*	139
5.7.1	Intervallo di ricerca	140
5.7.2	Stima iniziale	140
5.7.3	Tecniche di interpolazione	142
5.7.4	Criteri di arresto e fallimenti	149
5.8	Esercizi	150
6	Metodo del gradiente	151
6.1	Generalità	151
6.2	Definizione del metodo e proprietà di convergenza	152
6.3	Metodo del gradiente con passo costante	155
6.4	Rapidità di convergenza	156
6.5	Convergenza finita nel caso quadratico	159
6.6	Cenni sul metodo "Heavy Ball"	161
6.7	Esercizi	162
7	Metodo di Newton	165
7.1	Generalità	165
7.2	Convergenza locale	166

7.3	Metodo di Shamanskii	170
7.4	Globalizzazione del metodo di Newton	173
	7.4.1 Classificazione delle tecniche di globalizzazione	173
	7.4.2 Accettazione del passo unitario	175
	7.4.3 Condizioni sulla direzione di ricerca	177
7.5	Metodi ibridi	179
7.6	Modifiche della matrice Hessiana	182
7.7	Metodi di stabilizzazione non monotoni*	186
	7.7.1 Motivazioni	186
	7.7.2 Globalizzazione con ricerca unidimensionale non monotona	187
	7.7.3 Globalizzazione con strategia di tipo watchdog non monotona e ricerca unidimensionale non monotona	190
7.8	Convergenza a punti stazionari del "secondo ordine"*	197
	7.8.1 Concetti generali	197
	7.8.2 Metodo di ricerca unidimensionale curvilinea	198
	7.8.3 Proprietà sulle direzioni e analisi di convergenza	202
7.9	Esercizi	204
8	Metodi delle direzioni coniugate	207
8.1	Generalità	207
8.2	Definizioni e risultati preliminari	208
8.3	Metodo del gradiente coniugato: caso quadratico	213
	8.3.1 Il caso di matrice Hessiana semidefinita positiva*	217
	8.3.2 Minimi quadrati lineari	219
	8.3.3 Rapidità di convergenza	220
	8.3.4 Precondizionamento	227
8.4	Gradiente coniugato nel caso non quadratico	229
	8.4.1 Generalità e schema concettuale del metodo	229
	8.4.2 Metodo di Fletcher-Reeves*	233
	8.4.3 Metodo di Polyak-Polak-Ribière*	241
8.5	Esercizi	254
9	Metodi di trust region	255
9.1	Generalità	255
9.2	Il sufficiente decremento del modello quadratico e il passo di Cauchy	258
9.3	Analisi di convergenza globale*	260
9.4	Metodi di soluzione del sottoproblema	265
	9.4.1 Classificazione	265
	9.4.2 Condizioni necessarie e sufficienti di ottimalità	266
	9.4.3 Cenni sul calcolo della soluzione esatta*	269
	9.4.4 Metodo dogleg per il calcolo di una soluzione approssimata*	272

9.4.5	Metodo del gradiente coniugato di Steihaug per il calcolo di una soluzione approssimata*	275
9.5	Modifiche globalmente convergenti del metodo di Newton	278
9.6	Convergenza a punti stazionari del "secondo ordine"*	281
9.7	Esercizi	286
10	Metodi Quasi-Newton	289
10.1	Generalità	289
10.2	Formule di rango 1	292
10.3	Formule di rango 2	294
10.4	Convergenza globale metodo BFGS: caso convesso*	298
10.5	Condizioni di convergenza superlineare*	304
10.6	Rapidità di convergenza del metodo BFGS*	311
10.7	Esercizi	323
11	Metodo del gradiente di Barzilai-Borwein	325
11.1	Generalità	325
11.2	Metodo BB nel caso quadratico	326
11.3	Convergenza nel caso quadratico*	330
11.4	Estensioni del metodo BB*	337
11.5	Estensioni del metodo BB al caso non quadratico	338
11.6	Globalizzazione non monotona del metodo BB*	340
11.7	Esercizi	342
12	Metodi per problemi di minimi quadrati	343
12.1	Generalità	343
12.2	Problemi di minimi quadrati lineari	345
12.3	Metodi per problemi di minimi quadrati non lineari	348
12.3.1	Motivazioni	348
12.3.2	Metodo di Gauss-Newton	350
12.3.3	Metodo di Levenberg-Marquardt	357
12.4	Metodi incrementali: filtro di Kalman	364
12.5	Cenni sui metodi incrementali per problemi non lineari	366
12.6	Esercizi	368
13	Metodi per problemi a larga scala	371
13.1	Generalità	371
13.2	Metodo di Newton inesatto	372
13.3	Metodi di Newton troncato	376
13.3.1	Concetti generali	376
13.3.2	Metodo di Newton troncato basato su ricerca unidimensionale*	377
13.3.3	Metodo di Newton troncato di tipo trust region*	383
13.4	Metodi Quasi-Newton per problemi a larga scala	386
13.4.1	Concetti preliminari	386

13.4.2	Metodi Quasi-Newton senza memoria	386
13.4.3	Metodi Quasi-Newton a memoria limitata	387
13.5	Esercizi	391
14	Metodi senza derivate	393
14.1	Generalità	393
14.2	Metodi basati sull'approssimazione alle differenze finite	394
14.3	Metodo di Nelder-Mead	395
14.4	Metodi delle direzioni coordinate	398
14.4.1	Concetti preliminari	398
14.4.2	Metodo delle coordinate con semplice decremento	400
14.4.3	Una variante del metodo delle coordinate: metodo di Hooke-Jeeves	404
14.4.4	Metodi delle coordinate con sufficiente decremento	406
14.5	Metodi basati su direzioni che formano basi positive	412
14.6	Metodo delle direzioni coniugate	415
14.7	Cenni sui metodi basati su modelli di interpolazione	419
14.8	Cenni sul metodo "implicit filtering"	420
14.9	Esercizi	421
15	Metodi per sistemi di equazioni non lineari	423
15.1	Generalità	423
15.2	Metodi di tipo Newton	425
15.2.1	Globalizzazione di metodi di tipo Newton	426
15.3	Metodo di Broyden	430
15.4	Metodi basati sul residuo	432
15.5	Esercizi	439
16	Metodi di decomposizione	441
16.1	Generalità	441
16.2	Notazioni e tipi di decomposizione	444
16.3	Metodo di Gauss-Seidel a blocchi ed estensioni	446
16.3.1	Lo schema	446
16.3.2	Analisi di convergenza*	447
16.3.3	Modifiche del metodo di Gauss-Seidel	454
16.4	Metodi di discesa a blocchi	456
16.5	Metodo di Gauss-Southwell	458
16.6	Decomposizione con sovrapposizione dei blocchi	460
16.7	Metodo di Jacobi	464
16.8	Esercizi	466
17	Metodi per problemi con insieme ammissibile convesso	469
17.1	Generalità	469
17.2	Problemi con insieme ammissibile convesso	470
17.2.1	Direzioni ammissibili	470

17.2.2	Condizioni di ottimo con insieme ammissibile convesso	472
17.2.3	Problemi con vincoli lineari	474
17.2.4	Proiezione su un insieme convesso e condizioni di ottimo	478
17.3	Ricerca lungo una direzione ammissibile	483
17.4	Metodo di Frank-Wolfe (Conditional gradient method)	486
17.5	Metodo del gradiente proiettato	489
17.6	Convessità generalizzata: punti di minimo*	492
17.7	Esercizi	497
Appendice A	Richiami e notazioni	499
A.1	Lo spazio R^n come spazio lineare	499
A.2	Matrici e sistemi di equazioni lineari	501
A.3	Norma, metrica, topologia, prodotto scalare su R^n	504
A.4	Richiami e notazioni sulle matrici reali	510
A.5	Forme quadratiche	516
Appendice B	Richiami sulla differenziazione	519
B.1	Derivate del primo ordine di una funzione reale	519
B.2	Differenziazione di un vettore di funzioni	521
B.3	Derivate del secondo ordine di una funzione reale	523
B.4	Teorema della media e formula di Taylor	525
B.5	Derivazione di funzioni composte	527
B.6	Esempi	528
Appendice C	Convessità	535
C.1	Insiemi convessi	535
C.2	Funzioni convesse	544
C.3	Composizione di funzioni convesse	547
C.4	Proprietà di continuità delle funzioni convesse	550
C.5	Convessità di funzioni differenziabili	552
C.6	Monotonicità	556
C.7	Cenni sulla convessità generalizzata	558
Appendice D	Condizioni di ottimo per problemi vincolati	565
D.1	Condizioni di Fritz John	565
D.2	Qualificazione dei vincoli e condizioni di KKT	571
D.3	Moltiplicatori di Lagrange	576
D.4	Condizioni sufficienti nel caso convesso	577
D.5	Problemi con vincoli lineari	579
D.5.1	Problemi con vincoli di non negatività	580
D.5.2	Problemi con vincoli di "box"	581
D.5.3	Problemi con vincoli di simpleso	582
D.5.4	Programmazione quadratica	583
D.5.5	Programmazione lineare	584

Appendice E Aspetti numerici	587
E.1 Numeri in virgola mobile a precisione finita	587
E.2 Scala delle variabili e dell'obiettivo	589
E.3 Criteri di arresto e fallimenti	591
E.4 Differenze finite per l'approssimazione delle derivate	592
E.5 Cenni di differenziazione automatica	595
E.5.1 Il grafo computazionale	595
E.5.2 Il modo "diretto" di differenziazione automatica	597
E.5.3 Il modo "inverso" di differenziazione automatica	599
E.6 Alcuni problemi test di ottimizzazione non vincolata	601
Bibliografia	603
Indice analitico	611