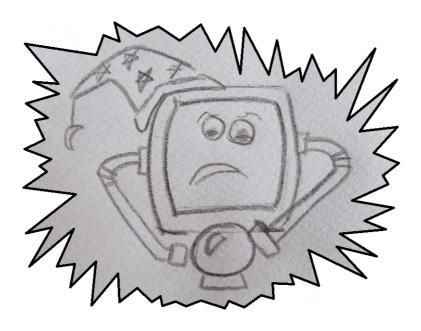
Fabio Marcuzzi

Analisi dei Dati mediante Modelli Matematici

Metodi numerici ed applicazioni a problemi di identificazione di modelli dai dati sperimentali, misura indiretta, predizione e stima dello stato

- 1^a edizione online -



Prefazione

In questo libro tratteremo un insieme fondamentale di metodi numerici per l'analisi dei dati, che fanno uso esplicito di modelli matematici esplicativi del fenomeno che ha generato i dati. In generale, i metodi quantitativi per l'analisi di dati numerici sono algoritmi di calcolo numerico, implementati per essere eseguiti dal processore adatto all'applicazione specifica: dalla workstation al microcontrollore embedded.

Negli ultimi tempi, l'elaborazione numerica dei dati è diventata pervasiva nelle attività quotidiane, sia professionali che personali (consumer). Essa è presente in vari tipi di applicazioni, effettuate da microprocessori residenti: in computers utilizzati per progettare beni materiali e strumenti di ogni genere, per effettuare analisi di mercato o test di laboratorio; nei sistemi elettronici di bordo negli aerei, nelle automobili, nelle apparecchiature domestiche, nella telefonia, nei motori di ricerca in internet, nei video-giochi ecc.

In generale, la presenza di un modello matematico di supporto all'algoritmo di elaborazione numerica permette a quest'ultimo di dare un risultato molto più aderente alla situazione specifica in cui esso sta operando. In questo modo si rendono possibili operazioni di predizione, stima, misura indiretta, ovvero la creazione di informazioni specifiche e non direttamente presenti nei dati di partenza.

In questo libro vengono presentati i metodi numerici fondamentali per l'analisi delle sequenze di dati nel dominio del tempo e della frequenza (e cenni ai metodi per l'analisi multi-risoluzione), per la soluzione di problemi di stima (usualmente ai minimi quadrati) di parametri lineari e nonlineari, per l'identificazione di modelli di ordine ridotto da collezioni di dati sperimentali e per la riduzione di modelli di dettaglio per via algebrica, ed un'introduzione ai metodi di regolarizzazione di problemi fortemente mal-condizionati o mal-posti, che sovente appaiono nella risoluzione numerica di problemi inversi. L'enfasi è posta, in particolare, sul fornire una certa visione dell'utilizzo di alcuni metodi numerici fondamentali in algoritmi più complessi di analisi dei dati e loro applicazioni. In particolare, viene mostrato come pochi metodi numerici fondamentali stiano alla base di molteplici e disparate applicazioni. Il materiale del libro è utilizzato nelle lezioni del corso di Metodi Numerici per l'Analisi dei Dati, tenuto

dall'autore ed inserito nel Corso di Laurea Magistrale in Matematica dell'Università di Padova, ma seguito anche da un discreto numero di studenti e dottorandi di ingegneria e delle scienze applicate. Il libro è rivolto a questo contesto multi-disciplinare, che è una parte importante della matematica applicata.

Il libro contiene una breve trattazione teorica, la descrizione in dettaglio degli algoritmi numerici, esempi ed implementazioni dei metodi numerici presentati ed alcune applicazioni, prevalentemente di problemi inversi nella meccanica computazionale, ma anche qualche esempio nel campo socio-economico. Il materiale è proposto in forma digitale sul web anche per quanto riguarda approfondimenti, esercizi, spiegazioni (es. mediante filmati), applicazioni ed implementazioni software. Il materiale sul web è ad accesso riservato al possessore di una copia del libro ed è frequentemente aggiornato: il lettore può quindi seguire i contenuti relativi ad una materia in costante e rapida evoluzione senza dover comprarsi necessariamente n edizioni successive del libro. Per accedere al materiale sul web è necessario registrarsi alla pagina web http://www.math.unipd.it/~marcuzzi/MNAD.html.

Pordenone e Padova, Settembre 2012

Fabio Marcuzzi
marcuzzi@math.unipd.it

L'autore intende devolvere i diritti derivanti dalla commercializzazione del presente libro alla cooperativa sociale ARCA (Allevamento Ricerca con Animali, http://www.arca.coop/): è una cooperativa di tibo B che è rivolta alla comunità allargata ed offre servizi di Pet Teraphy ed Onoterapia alle persone svantaggiate (bambini in situazioni di disagio, anziani, giovani dei centri per i disturbi alimentari, ecc.). Inoltre, il personale dipendente della cooperativa, addetto all'addestramento degli animali (cani razza Labrador ed asinelli) è diversamente abile ed a loro, la cooperativa, offre una opportunità di lavoro responsabile che li rende realmente autonomi (attualmente ARCA ha quattro dipendenti, tre sono portatori di handicap). Il suo buon funzionamento è garantito dal volontariato creatosi attorno. Per saperne di più: http://www.arca.coop/

L'immagine di copertina è stata realizzata da un bambino e raffigura scherzosamente le previsioni al computer come una magia, ma è stata scelta per il senso di creatività che in realtà questo disegno ispira (il mago sembra più un artista modellatore). La creatività è un ingrediente fondamentale per la matematica applicata, sia alle scienze che all'ingegneria. E' attribuita al matematico Henri Poincaré la frase: "Creatività è unire elementi esistenti con connessioni nuove, che siano utili".

Indice

P	reraz	ione		U
Ι	\mathbf{Se}	gnali	e sistemi	1
1	Dat	i quan	titativi	5
	1.1	Origin	ne e proprieta' dei dati	5
	1.2	Seque	nze di Dati, Segnali a tempo discreto e Serie storiche	7
	1.3	Camp	ionamento di segnali continui	9
	1.4	Aleato	prietá nei Dati	10
2	Mo	delli m	natematici	13
	2.1	Model	lli di regressione (statici)	14
		2.1.1	Modello lineare standard	14
		2.1.2	Modello lineare generalizzato (Gauss-Markov)	15
		2.1.3	Modello lineare con errori nelle variabili	15
		2.1.4	Modello non-lineare perceptron	16
		2.1.5	Modello non-lineare $\textit{multi-layer neural network}$	16
		2.1.6	Modello non-lineare kernel-based	18
	2.2	Model	lli di auto-regressione (sistemi dinamici)	19
		2.2.1	Modello lineare DLTI e convoluzione discreta	19
		2.2.2	Modello lineare ARMA ed equazioni alle differenze	21
			2.2.2.1 Modello ARMA stocastico	24
		2.2.3	Modello lineare con variabili di stato ($state\text{-}space\ model)$	24
			2.2.3.1 Modello state-space stocastico	27
		2.2.4	Modello non-lineare kernel-based con variabili di stato	27

3	Ana	disi qu	antitative	29
	3.1	Filtrag	ggio digitale	29
	3.2	Previs	ione	29
	3.3	Identii	ficazione e stima di parametri	30
	3.4	Invers	ione o deconvoluzione	31
II	\mathbf{N}	Ietodi	i Numerici	33
4	Alg	ebra li	neare numerica	37
	4.1	La fat	torizzazione QR	37
		4.1.1	Metodo di Gram-Schmidt	38
		4.1.2	Metodo di Householder	38
		4.1.3	Analisi all'indietro dell'errore	40
		4.1.4	Formulazione ricorsiva della fattorizzazione QR	41
		4.1.5	Trasformazioni di Givens	42
	4.2	La Sin	agular Value Decomposition (SVD)	43
		4.2.1	Intepretazione geometrica della SVD	44
		4.2.2	Rilevanza della SVD nel calcolo numerico	44
		4.2.3	Calcolo della SVD	46
		4.2.4	Pseudo-inversa di una matrice	48
		4.2.5	Analisi di sistemi lineari con la SVD	49
	4.3	Metod	li numerici per il problema dei minimi quadrati lineari standard	50
		4.3.1	Metodo delle Equazioni Normali	51
		4.3.2	Caratterizzazione della soluzione ai minimi quadrati	51
		4.3.3	Metodo con la Fattorizzazione QR $\ \ldots \ \ldots \ \ldots$	52
		4.3.4	Metodo con la SVD	52
		4.3.5	Scelta dell'algoritmo più adatto	52
			4.3.5.1 Effetto della presenza di rumore nei dati	53
	4.4		li numerici per il problema dei minimi quadrati totali (Total Least es)	54

5	Reg	golarizzazione di problemi mal-condizionati o mal-posti	57
	5.1	Analisi ai valori singolari del malcondizionamento di una matrice ${\cal A}$	58
	5.2	La SVD Troncata	59
		5.2.1 Soluzione di problemi ai minimi quadrati mal-condizionati o	
		mal-posti	59
		5.2.2 Scelta dell'indice di troncamento	61
	5.3	Il metodo di Tikhonov	62
	5.4	Il metodo iterativo di Landweber	63
6	Rid	uzione d'ordine del modello	65
	6.1	Riduzione di modelli lineari statici: PCA	65
	6.2	Riduzione di modelli DLTI nella rappresentazione state-space	66
		6.2.1 Troncamento previo bilanciamento	67
		6.2.2 Troncamento modale	68
	6.3	Proper Orthogonal Decomposition (POD) $\ \ldots \ \ldots \ \ldots \ \ldots$	69
7	Met	todi numerici per il problema dei minimi quadrati nonlineari	73
	7.1	Metodo di Newton	74
	7.2	Metodo di Gauss-Newton	76
	7.3	Il metodo di Levenberg-Marquardt	76
		7.3.1 Scelta di $\lambda^{(k)}$	77
		7.3.2 Scelta di $\Delta^{(k)}$	77
Η	\mathbf{I} A	Analisi mediante modelli non parametrici	7 9
8	Ana	alisi nel dominio dei dati	83
	8.1	Correlazione e Covarianza di Segnali Discreti	83
		8.1.1 Aleatorietà nei Dati	84
		8.1.2 Stima dell'aspettazione	84
	8.2	Principal Component Analysis (PCA)	85
		8.2.1 Calcolo delle componenti principali	85
	8.3	Canonical Correlation Analysis (CCA)	86
	8.4	Kernel PCA	88
		8.4.1 Formulazione duale della PCA	88
	8.5	Compressed sensing	89

9	Ana	disi di	Fourier	nel discreto	91
	9.1	La Tra	sformata	di Fourier Discreta (DFT)	91
	9.2	Rappr	esentazio	ne dei Segnali Discreti mediante DFT	93
		9.2.1	Sequenz	e di Dati Periodiche di Durata Illimitata	93
		9.2.2	Sequenz	e di Dati Aperiodiche di Durata Illimitata	94
		9.2.3	Sequenz	e di Dati di Durata Limitata	95
		9.2.4	Sequenz	e di Dati ad Energia Infinita	96
	9.3	Rappr	esentazio	ne dei Sistemi DLTI mediante DFT	96
		9.3.1	Risposta	a in Frequenza di un Sistema DLTI	96
		9.3.2	Trasform	nata z e Funzione di Trasferimento	98
	9.4	Calcol	o della D	FT: algoritmo della Fast Fourier Transform (FFT) $$	98
		9.4.1	Algoritn	no della FFT per segnali mono-dimensionali	98
			9.4.1.1	Una semplice riduzione del numero di operazioni	100
			9.4.1.2	Il calcolo sul posto	100
			9.4.1.3	La gestione degli indici	101
		9.4.2	Utilizzo	in pratica della DFT	101
			9.4.2.1	Rappresentazione grafica della DFT	101
			9.4.2.2	Il problema del $leakage$ e la finestratura dei dati $\ .\ .\ .$	102
			9.4.2.3	il problema dell'effetto <i>picket fence</i> e la tecnica di <i>zero-padding</i>	103
			9.4.2.4	Risoluzione in frequenza	103
			9.4.2.5	Presenza di rumore nei dati	104
		9.4.3		ello Spettro di Potenza di un Segnale	104
		9.4.4		e l'algoritmo FFT per segnali bi-dimensionali	106
10	O Ana	disi ter	npo-free	quenza	109
	10.1	Short-	Time Fou	rrier Transform (STFT)	109
	10.2	La Tra	sformata	Wavelet	110
		10.2.1	Le wave	lets di Haar	111
				lets della Daubechies [?]	112
		10.2.3	Le Coifl	ets	113
		10.2.4	Conserv	azione e compattamento dell'energia	113
		10.2.5	Interpre	tazione della trasformata wavelet come banco di filtri [?]	113
		10.2.6	Analisi i	n frequenza delle wavelets	114

IV	A	analis	i mediante modelli parametrici	117
11	Iden	ntificaz	zione di modelli lineari	121
	11.1	Identif	icazione di modelli lineari statici	121
		11.1.1	Stima dei parametri del modello lineare standard	121
		11.1.2	Stima dei parametri del modello lineare generalizzato	122
		11.1.3	Stima di parametri tempo-varianti	123
		11.1.4	Determinazione dell'ordine del modello dai dati sperimentali .	123
	11.2	Identif	ficazione di sistemi DLTI da dati privi di rumore	124
		11.2.1	Caso di risposta al campione unitario e rappresentazione ARMA	125
		11.2.2	Caso di risposta al campione unitario e rappresentazione $State-Space$	127
		11.2.3	Eccitazione persistente	127
		11.2.4	Caso di coppie ingresso/uscita qualsiasi e rappresentazione ARMA	128
		11.2.5	Caso di coppie ingresso/uscita qualsiasi e rappresentazione $state\text{-}space$	129
-	11.3	Identif	ficazione di modelli DLTI da dati con rumore	130
		11.3.1	Analisi dei valori singolari e presenza di rumore nei dati $\ . \ . \ .$	130
		11.3.2	Determinazione dell'ordine di un modello ARX $\ .\ .\ .\ .\ .$	130
			11.3.2.1 L'analisi dei valori singolari	131
			11.3.2.2 Criteri di parsimonia	131
-	11.4	Analis	i parametrica di <i>serie storiche</i>	131
		11.4.1	Scomposizione della serie storica	132
		11.4.2	Predizione lineare	133
1 2]	Iden	ntificaz	ione di modelli nonlineari	135
			di parametri nonlineari in modelli DLTI	135
			Presenza di rumore nelle equazioni del modello e nei dati sperimentali	137
			12.1.1.1 Calcolo di $\psi(\theta)$	137
			12.1.1.2 Significatività dell'aggiornamento delle stime	138
		12.1.2	Malcondizionamento e regolarizzazione delle stime	138
	12.2		ndimento di una rete neurale	
			Il perceptron e la separazione mediante iperpiani	

		12.2.2 L'algoritmo di back-propagation	140
		12.2.3 Procedura di apprendimento	143
		12.2.4 Momentum Algorithm	144
	12.3	Metodi kernel	145
		12.3.1 Identificazione di modelli nonlineari statici	145
		12.3.2 Support Vector Machines	146
13	Stin	na dello stato di sistemi dinamici 1	47
	13.1	Stima ricorsiva dello stato di modelli DLTI nella rappresentazione $\it state$	
		7	147
		13.1.1 costruzione dell'osservatore	148
	13.2	Stima dello stato in presenza di rumore di modello e di misura: il Filtro di Kalman	149
			149
			151
		13.2.3 Implementazioni Square-Root del filtro di Kalman	151
	13.3	Stima dello stato iniziale in sistemi evolutivi nonlineari	152
		13.3.1 Discretizzazione nel tempo	153
		13.3.2 Stima delle condizioni iniziali (stato iniziale)	154
		13.3.3 Il metodo del modello aggiunto	155
		13.3.4 Procedura di ottimizzazione con il modello aggiunto	157
	_		
\mathbf{V}	Pı	roblemi inversi ed applicazioni 1	59
14	Pro	blemi inversi in meccanica computazionale	63
	14.1	Scelta del modello numerico	163
	14.2	Modelli a parametri concentrati	164
	14.3	Modelli a parametri distribuiti	164
		14.3.1Relazioni tra modello geometrico ed algebrico del sistema	164
	14.4	Rappresentazione in forma nello spazio degli stati ($state\text{-}space$)	164
	14.5	Modelli misti a parametri concentrati e distribuiti	165
	14.6	Il controllo non-distruttivo dei materiali	165
		14.6.1 Scelta del metodo numerico di analisi dei dati	166
	14 7	Model-based fault-detection per i sistemi di controllo fault-tolerant	167

15	App	olicazio	oni	169
	15.1	Stima	della corrosione mediante termografie	169
	15.2	Essica	zione a controllo temperometrico	169
	15.3	Stima	delle forze di contatto tra due corpi in movimento relativo $\ . \ . \ .$	170
	15.4	Contro	ollo adattativo di un sistema a corpi rigidi	170
	15.5	Stima	del peso di corpi in rotazione	170
	15.6	Tomog	grafia a vibrazioni elastiche: la geosonda	171
	15.7	Model	llazione state-space di dati relativi al mercato immobiliare	171
\mathbf{A}	Ana	disi de	ei sistemi DLTI state-space	173
	A.1	Analis	si modale	174
	A.2	Raggiu	ungibilità	175
	A.3	Osserv	vabilità	175
В	Solu	ızione	numerica di problemi differenziali evolutivi	179
	B.1	Fenom	neni diffusivi	179
	B.2	Fenom	neni di trasporto e propagazione di onde	181
	В.3		one numerica delle equazioni della dinamica di un sistema nico elastico lineare	181
		B.3.1	Integrazione nel tempo	181
		B.3.2	Metodi per sistemi di equazioni differenziali del primo ordine	
			nel tempo	182
			B.3.2.1 Eulero esplicito (forward Euler, $\theta = 0$)	183
			B.3.2.2 Eulero implicito (backward Euler, $\theta = 1$)	183
			B.3.2.3 Crank-Nicholson $(\theta = \frac{1}{2})$	183
\mathbf{C}	Ling	guaggi	di programmazione per il calcolo numerico	185
	C.1	Criter	i di scelta del linguaggio	185
	C.2	Introd	luzione a Python	186
		C.2.1	Sviluppo di programmi in Python	187
		C.2.2	Documentazione su Python	187