

Fundamentals of Speaker Recognition

Homayoon Beigi

Fundamentals of Speaker Recognition

 Springer

Dr. Homayoon Beigi
Recognition Technologies, Inc.
Yorktown Heights New York, NY
USA
beigi@recotechnologies.com

ISBN 978-0-387-77591-3 e-ISBN 978-0-387-77592-0
DOI 10.1007/978-0-387-77592-0
Springer New York Dordrecht Heidelberg London

Library of Congress Control Number: 2011941119

© Springer Science+Business Media, LLC 2012

All rights reserved. This work may not be translated or copied in whole or in part without the written permission of the publisher (Springer Science+Business Media, LLC, 233 Spring Street, New York, NY 10013, USA), except for brief excerpts in connection with reviews or scholarly analysis. Use in connection with any form of information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed is forbidden.

The use in this publication of trade names, trademarks, service marks, and similar terms, even if they are not identified as such, is not to be taken as an expression of opinion as to whether or not they are subject to proprietary rights.

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

Contents

Part I Basic Theory

1	Introduction	3
1.1	Definition and History	3
1.2	Speaker Recognition Branches	5
1.2.1	Speaker Verification (Speaker Authentication)	5
1.2.2	Speaker Identification (Closed-Set and Open-Set)	7
1.2.3	Speaker and Event Classification	8
1.2.4	Speaker Segmentation	9
1.2.5	Speaker Detection	11
1.2.6	Speaker Tracking	11
1.3	Speaker Recognition Modalities	12
1.3.1	Text-Dependent Speaker Recognition	12
1.3.2	Text-Independent Speaker Recognition	13
1.3.3	Text-Prompted Speaker Recognition	14
1.3.4	Knowledge-Based Speaker Recognition	15
1.4	Applications	16
1.4.1	Financial Applications	16
1.4.2	Forensic and Legal Applications	18
1.4.3	Access Control (Security) Applications	19
1.4.4	Audio and Video Indexing (Diarization) Applications	19
1.4.5	Surveillance Applications	20
1.4.6	Teleconferencing Applications	21
1.4.7	Proctorless Oral Testing	21
1.4.8	Other Applications	23
1.5	Comparison to Other Biometrics	23
1.5.1	Deoxyribonucleic Acid (DNA)	24
1.5.2	Ear	25
1.5.3	Face	27
1.5.4	Fingerprint and Palm	28
1.5.5	Hand and Finger Geometry	30

1.5.6	Iris	30
1.5.7	Retina	31
1.5.8	Thermography	32
1.5.9	Vein	32
1.5.10	Gait	33
1.5.11	Handwriting	34
1.5.12	Keystroke	35
1.5.13	Multimodal	35
1.5.14	Summary of Speaker Biometric Characteristics	37
	References	38
2	The Anatomy of Speech	43
2.1	The Human Vocal System	44
2.1.1	Trachea and Larynx	44
2.1.2	Vocal Folds (Vocal Chords)	44
2.1.3	Pharynx	47
2.1.4	Soft Palate and the Nasal System	48
2.1.5	Hard Palate	48
2.1.6	Oral Cavity Exit	48
2.2	The Human Auditory System	48
2.2.1	The Ear	50
2.3	The Nervous System and the Brain	51
2.3.1	Neurons – Elementary Building Blocks	52
2.3.2	The Brain	54
2.3.3	Function Localization in the Brain	59
2.3.4	Specializations of the Hemispheres of the Brain	62
2.3.5	Audio Production	64
2.3.6	Auditory Perception	66
2.3.7	Speaker Recognition	71
	References	72
3	Signal Representation of Speech	75
3.1	Sampling The Audio	77
3.1.1	The Sampling Theorem	78
3.1.2	Convergence Criteria for the Sampling Theorem	84
3.1.3	Extensions of the Sampling Theorem	84
3.2	Quantization and Amplitude Errors	85
3.3	The Speech Waveform	87
3.4	The Spectrogram	87
3.5	Formant Representation	89
3.6	Practical Sampling and Associated Errors	92
3.6.1	Ideal Sampler	98
3.6.2	Aliasing	99
3.6.3	Truncation Error	102
3.6.4	Jitter	103

3.6.5	Loss of Information	104
References	105
4	Phonetics and Phonology	107
4.1	Phonetics	107
4.1.1	Initiation	109
4.1.2	Phonation	109
4.1.3	Articulation	110
4.1.4	Coordination	111
4.1.5	Vowels	112
4.1.6	Pulmonic Consonants	115
4.1.7	Whisper	119
4.1.8	Whistle	119
4.1.9	Non-Pulmonic Consonants	120
4.2	Phonology and Linguistics	122
4.2.1	Phonemic Utilization Across Languages	122
4.2.2	Whisper	125
4.2.3	Importance of Vowels in Speaker Recognition	127
4.2.4	Evolution of Languages toward Discriminability	129
4.3	Suprasegmental Features of Speech	131
4.3.1	Prosodic Features	132
4.3.2	Metrical features of Speech	138
4.3.3	Temporal features of Speech	140
4.3.4	Co-Articulation	140
References	141
5	Signal Processing of Speech and Feature Extraction	143
5.1	Auditory Perception	144
5.1.1	Pitch	146
5.1.2	Loudness	149
5.1.3	Timbre	151
5.2	The Sampling Process	152
5.2.1	Anti-Aliasing	153
5.2.2	Hi-Pass Filtering	153
5.2.3	Pre-Emphasis	153
5.2.4	Quantization	155
5.3	Spectral Analysis and Direct Method Features	157
5.3.1	Framing the Signal	160
5.3.2	Windowing	162
5.3.3	Discrete Fourier Transform (DFT) and Spectral Estimation	167
5.3.4	Frequency Warping	169
5.3.5	Magnitude Warping	172
5.3.6	Mel Frequency Cepstral Coefficients (MFCC)	173
5.3.7	Mel Cepstral Dynamics	175
5.4	Linear Predictive Cepstral Coefficients (LPCC)	176

5.4.1	Autoregressive (AR) Estimate of the PSD	177
5.4.2	LPC Computation	184
5.4.3	Partial Correlation (PARCOR) Features	185
5.4.4	Log Area Ratio (LAR) Features	189
5.4.5	Linear Predictive Cepstral Coefficient (LPCC) Features	189
5.5	Perceptual Linear Predictive (PLP) Analysis	190
5.5.1	Spectral Analysis	191
5.5.2	Bark Frequency Warping	191
5.5.3	Equal-Loudness Pre-emphasis	192
5.5.4	Magnitude Warping	193
5.5.5	Inverse DFT	193
5.6	Other Features	193
5.6.1	Wavelet Filterbanks	194
5.6.2	Instantaneous Frequencies	197
5.6.3	Empirical Mode Decomposition (EMD)	198
5.7	Signal Enhancement and Pre-Processing	199
	References	199
6	Probability Theory and Statistics	205
6.1	Set Theory	205
6.1.1	Equivalence and Partitions	208
6.1.2	R-Rough Sets (Rough Sets)	210
6.1.3	Fuzzy Sets	211
6.2	Measure Theory	211
6.2.1	Measure	212
6.2.2	Multiple Dimensional Spaces	216
6.2.3	Metric Space	217
6.2.4	Banach Space (Normed Vector Space)	218
6.2.5	Inner Product Space (Dot Product Space)	219
6.2.6	Infinite Dimensional Spaces (Pre-Hilbert and Hilbert)	219
6.3	Probability Measure	221
6.4	Integration	227
6.5	Functions	228
6.5.1	Probability Density Function	229
6.5.2	Densities in the Cartesian Product Space	232
6.5.3	Cumulative Distribution Function	235
6.5.4	Function Spaces	236
6.5.5	Transformations	238
6.6	Statistical Moments	239
6.6.1	Mean	239
6.6.2	Variance	242
6.6.3	Skewness (skew)	245
6.6.4	Kurtosis	246
6.7	Discrete Random Variables	247
6.7.1	Combinations of Random Variables	250

6.7.2	Convergence of a Sequence	250
6.8	Sufficient Statistics	251
6.9	Moment Estimation	253
6.9.1	Estimating the Mean	253
6.9.2	Law of Large Numbers (LLN)	254
6.9.3	Different Types of Mean	257
6.9.4	Estimating the Variance	258
6.10	Multi-Variate Normal Distribution	259
	References	261
7	Information Theory	265
7.1	Sources	266
7.2	The Relation between Uncertainty and Choice	269
7.3	Discrete Sources	269
7.3.1	Entropy or Uncertainty	270
7.3.2	Generalized Entropy	278
7.3.3	Information	279
7.3.4	The Relation between Information and Entropy	280
7.4	Discrete Channels	282
7.5	Continuous Sources	284
7.5.1	Differential Entropy (Continuous Entropy)	284
7.6	Relative Entropy	286
7.6.1	Mutual Information	291
7.7	Fisher Information	294
	References	299
8	Metrics and Divergences	301
8.1	Distance (Metric)	301
8.1.1	Distance Between Sequences	302
8.1.2	Distance Between Vectors and Sets of Vectors	302
8.1.3	Hellinger Distance	304
8.2	Divergences and Directed Divergences	304
8.2.1	Kullback-Leibler's Directed Divergence	305
8.2.2	Jeffreys' Divergence	305
8.2.3	Bhattacharyya Divergence	306
8.2.4	Matsushita Divergence	307
8.2.5	F-Divergence	308
8.2.6	δ -Divergence	309
8.2.7	χ^α Directed Divergence	310
	References	310
9	Decision Theory	313
9.1	Hypothesis Testing	313
9.2	Bayesian Decision Theory	316
9.2.1	Binary Hypothesis	320

9.2.2	Relative Information and Log Likelihood Ratio	321
9.3	Bayesian Classifier	322
9.3.1	Multi-Dimensional Normal Classification	326
9.3.2	Classification of a Sequence	328
9.4	Decision Trees	331
9.4.1	Tree Construction	332
9.4.2	Types of Questions	333
9.4.3	Maximum Likelihood Estimation (MLE)	336
	References	338
10	Parameter Estimation	341
10.1	Maximum Likelihood Estimation	342
10.2	Maximum A-Posteriori (MAP) Estimation	344
10.3	Maximum Entropy Estimation	345
10.4	Minimum Relative Entropy Estimation	346
10.5	Maximum Mutual Information Estimation (MMIE)	348
10.6	Model Selection	349
10.6.1	Akaike Information Criterion (AIC)	350
10.6.2	Bayesian Information Criterion (BIC)	353
	References	354
11	Unsupervised Clustering and Learning	357
11.1	Vector Quantization (VQ)	358
11.2	Basic Clustering Techniques	359
11.2.1	Standard k-Means (Lloyd) Algorithm	360
11.2.2	Generalized Clustering	363
11.2.3	Overpartitioning	364
11.2.4	Merging	364
11.2.5	Modifications to the k-Means Algorithm	365
11.2.6	k-Means Wrappers	368
11.2.7	Rough k-Means	375
11.2.8	Fuzzy k-Means	377
11.2.9	k-Harmonic Means Algorithm	378
11.2.10	Hybrid Clustering Algorithms	380
11.3	Estimation using Incomplete Data	381
11.3.1	Expectation Maximization (EM)	381
11.4	Hierarchical Clustering	388
11.4.1	Agglomerative (Bottom-Up) Clustering (AHC)	389
11.4.2	Divisive (Top-Down) Clustering (DHC)	389
11.5	Semi-Supervised Learning	390
	References	390

12 Transformation	393
12.1 Principal Component Analysis (PCA)	394
12.1.1 Formulation	394
12.2 Generalized Eigenvalue Problem	397
12.3 Nonlinear Component Analysis	399
12.3.1 Kernel Principal Component Analysis (Kernel PCA)	400
12.4 Linear Discriminant Analysis (LDA)	401
12.4.1 Integrated Mel Linear Discriminant Analysis (IMELDA)	404
12.5 Factor Analysis	404
References	409
13 Hidden Markov Modeling (HMM)	411
13.1 Memoryless Models	413
13.2 Discrete Markov Chains	415
13.3 Markov Models	416
13.4 Hidden Markov Models	418
13.5 Model Design and States	421
13.6 Training and Decoding	423
13.6.1 Trellis Diagram Representation	428
13.6.2 Forward Pass Algorithm	430
13.6.3 Viterbi Algorithm	432
13.6.4 Baum-Welch (Forward-Backward) Algorithm	433
13.7 Gaussian Mixture Models (GMM)	442
13.7.1 Training	444
13.7.2 Tractability of Models	449
13.8 Practical Issues	451
13.8.1 Smoothing	451
13.8.2 Model Comparison	453
13.8.3 Held-Out Estimation	456
13.8.4 Deleted Estimation	461
References	462
14 Neural Networks	465
14.1 Perceptron	466
14.2 Feedforward Networks	466
14.2.1 Auto Associative Neural Networks (AANN)	469
14.2.2 Radial Basis Function Neural Networks (RBFNN)	469
14.2.3 Training (Learning) Formulation	470
14.2.4 Optimization Problem	473
14.2.5 Global Solution	474
14.3 Recurrent Neural Networks (RNN)	476
14.4 Time-Delay Neural Networks (TDNNs)	477
14.5 Hierarchical Mixtures of Experts (HME)	479
14.6 Practical Issues	479
References	481

15 Support Vector Machines	485
15.1 Risk Minimization	488
15.1.1 Empirical Risk Minimization	492
15.1.2 Capacity and Bounds on Risk	493
15.1.3 Structural Risk Minimization	493
15.2 The Two-Class Problem	494
15.2.1 Dual Representation	497
15.2.2 Soft Margin Classification	500
15.3 Kernel Mapping	503
15.3.1 The Kernel Trick	504
15.4 Positive Semi-Definite Kernels	506
15.4.1 Linear Kernel	506
15.4.2 Polynomial Kernel	506
15.4.3 Gaussian Radial Basis Function (GRBF) Kernel	507
15.4.4 Cosine Kernel	508
15.4.5 Fisher Kernel	508
15.4.6 GLDS Kernel	509
15.4.7 GMM-UBM Mean Interval (GUMI) Kernel	510
15.5 Non Positive Semi-Definite Kernels	511
15.5.1 Jeffreys Divergence Kernel	511
15.5.2 Fuzzy Hyperbolic Tangent (tanh) Kernel	512
15.5.3 Neural Network Kernel	513
15.6 Kernel Normalization	513
15.7 Kernel Principal Component Analysis (Kernel PCA)	514
15.8 Nuisance Attribute Projection (NAP)	516
15.9 The multiclass (Γ -Class) Problem	518
References	519

Part II Advanced Theory

16 Speaker Modeling	525
16.1 Individual Speaker Modeling	526
16.2 Background Models and Cohorts	527
16.2.1 Background Models	528
16.2.2 Cohorts	529
16.3 Pooling of Data and Speaker Independent Models	529
16.4 Speaker Adaptation	530
16.4.1 Factor Analysis (FA)	530
16.4.2 Joint Factor Analysis (JFA)	531
16.4.3 Total Factors (Total Variability)	532
16.5 Audio Segmentation	532
16.6 Model Quality Assessment	534
16.6.1 Enrollment Utterance Quality Control	534
16.6.2 Speaker Menagerie	536
References	538

17	Speaker Recognition	543
17.1	The Enrollment Task	543
17.2	The Verification Task	544
17.2.1	Text-Dependent	546
17.2.2	Text-Prompted	546
17.2.3	Knowledge-Based	548
17.3	The Identification Task	548
17.3.1	Closed-Set Identification	548
17.3.2	Open-Set Identification	549
17.4	Speaker Segmentation	549
17.5	Speaker and Event Classification	550
17.5.1	Gender and Age Classification (Identification)	551
17.5.2	Audio Classification	553
17.5.3	Multiple Codebooks	553
17.5.4	Farfield Speaker Recognition	553
17.5.5	Whispering Speaker Recognition	554
17.6	Speaker Diarization	554
17.6.1	Speaker Position and Orientation	555
	References	555
18	Signal Enhancement and Compensation	561
18.1	Silence Detection, Voice Activity Detection (VAD)	561
18.2	Audio Volume Estimation	564
18.3	Echo Cancellation	564
18.4	Spectral Filtering and Cepstral Liftering	565
18.4.1	Cepstral Mean Normalization (Subtraction) – CMN (CMS)	567
18.4.2	Cepstral Mean and Variance Normalization (CMVN)	569
18.4.3	Cepstral Histogram Normalization (Histogram Equalization)	570
18.4.4	RelAtive SpecTrAl (RASTA) Filtering	571
18.4.5	Other Lifters	571
18.4.6	Vocal Tract Length Normalization (VTLN)	573
18.4.7	Other Normalization Techniques	576
18.4.8	Steady Tone Removal (Narrowband Noise Reduction)	579
18.4.9	Adaptive Wiener Filtering	580
18.5	Speaker Model Normalization	581
18.5.1	Z-Norm	581
18.5.2	T-Norm (Test Norm)	582
18.5.3	H-Norm	582
18.5.4	HT-Norm	582
18.5.5	AT-Norm	582
18.5.6	C-Norm	582
18.5.7	D-Norm	583
18.5.8	F-Norm (F-Ratio Normalization)	583
18.5.9	Group-Specific Normalization	583

18.5.10 Within Class Covariance Normalization (WCCN)	583
18.5.11 Other Normalization Techniques	583
References	584

Part III Practice

19 Evaluation and Representation of Results	589
19.1 Verification Results	589
19.1.1 Equal-Error Rate	589
19.1.2 Half Total Error Rate	590
19.1.3 Receiver Operating Characteristic (ROC) Curve	590
19.1.4 Detection Error Trade-Off (DET) Curve	592
19.1.5 Detection Cost Function (DCF)	593
19.2 Identification Results	593
References	594
20 Time Lapse Effects (Case Study)	595
20.1 The Audio Data	598
20.2 Baseline Speaker Recognition	598
References	600
21 Adaptation over Time (Case Study)	601
21.1 Data Augmentation	601
21.2 Maximum A Posteriori (MAP) Adaptation	603
21.3 Eigenvoice Adaptation	605
21.4 Minimum Classification Error (MCE)	605
21.5 Linear Regression Techniques	606
21.5.1 Maximum Likelihood Linear Regression (MLLR)	606
21.6 Maximum a-Posteriori Linear Regression (MAPLR)	607
21.6.1 Other Adaptation Techniques	607
21.7 Practical Perspectives	607
References	608
22 Overall Design	611
22.1 Choosing the Model	611
22.1.1 Phonetic Speaker Recognition	612
22.2 Choosing an Adaptation Technique	613
22.3 Microphones	613
22.4 Channel Mismatch	615
22.5 Voice Over Internet Protocol (VoIP)	615
22.6 Public Databases	616
22.6.1 NIST	616
22.6.2 Linguistic Data Consortium (LDC)	616
22.6.3 European Language Resources Association (ELRA)	619
22.7 High Level Information	620
22.7.1 Choosing Basic Segments	622

22.8	Numerical Stability	623
22.9	Privacy	624
22.10	Biometric Encryption	625
22.11	Spoofing	625
	22.11.1 Text-Prompted Verification Systems	625
	22.11.2 Text-Independent Verification Systems	626
22.12	Quality Issues	627
22.13	Large-Scale Systems	628
22.14	Useful Tools	628
	References	629

Part IV Background Material

23	Linear Algebra	635
23.1	Basic Definitions	635
23.2	Norms	636
23.3	Gram-Schmidt Orthogonalization	641
	23.3.1 Ordinary Gram-Schmidt Orthogonalization	641
	23.3.2 Modified Gram-Schmidt Orthogonalization	641
23.4	Sherman-Morrison Inversion Formula	642
23.5	Vector Representation under a Set of Normal Conjugate Direction	642
23.6	Stochastic Matrix	643
23.7	Linear Equations	643
	References	646
24	Integral Transforms	647
24.1	Complex Variable Theory in Integral Transforms	648
	24.1.1 Complex Variables	648
	24.1.2 Limits	651
	24.1.3 Continuity and Forms of Discontinuity	652
	24.1.4 Convexity and Concavity of Functions	658
	24.1.5 Odd, Even and Periodic Functions	661
	24.1.6 Differentiation	663
	24.1.7 Analyticity	665
	24.1.8 Integration	672
	24.1.9 Power Series Expansion of Functions	683
	24.1.10 Residues	686
24.2	Relations Between Functions	688
	24.2.1 Convolution	688
	24.2.2 Correlation	689
24.3	Orthogonality of Functions	690
24.4	Integral Equations	694
24.5	Kernel Functions	696
	24.5.1 Hilbert's Expansion Theorem	698
	24.5.2 Eigenvalues and Eigenfunctions of the Kernel	700

24.6	Fourier Series Expansion	708
24.6.1	Convergence of the Fourier Series	713
24.6.2	Parseval's Theorem	714
24.7	Wavelet Series Expansion	716
24.8	The Laplace Transform	717
24.8.1	Inversion	720
24.8.2	Some Useful Transforms	721
24.9	Complex Fourier Transform (Fourier Integral Transform)	722
24.9.1	Translation	724
24.9.2	Scaling	724
24.9.3	Symmetry Table	724
24.9.4	Time and Complex Scaling and Shifting	725
24.9.5	Convolution	725
24.9.6	Correlation	726
24.9.7	Parseval's Theorem	726
24.9.8	Power Spectral Density	728
24.9.9	One-Sided Power Spectral Density	728
24.9.10	PSD-per-unit-time	729
24.9.11	Wiener-Khintchine Theorem	729
24.10	Discrete Fourier Transform (DFT)	731
24.10.1	Inverse Discrete Fourier Transform (IDFT)	732
24.10.2	Periodicity	734
24.10.3	Plancherel and Parseval's Theorem	734
24.10.4	Power Spectral Density (PSD) Estimation	735
24.10.5	Fast Fourier Transform (FFT)	736
24.11	Discrete-Time Fourier Transform (DTFT)	738
24.11.1	Power Spectral Density (PSD) Estimation	739
24.12	Complex Short-Time Fourier Transform (STFT)	740
24.12.1	Discrete-Time Short-Time Fourier Transform DTSTFT	744
24.12.2	Discrete Short-Time Fourier Transform DSTFT	746
24.13	Discrete Cosine Transform (DCT)	748
24.13.1	Efficient DCT Computation	749
24.14	The z -Transform	750
24.14.1	Translation	756
24.14.2	Scaling	756
24.14.3	Shifting – Time Lag	757
24.14.4	Shifting – Time Lead	757
24.14.5	Complex Translation	757
24.14.6	Initial Value Theorem	758
24.14.7	Final Value Theorem	758
24.14.8	Real Convolution Theorem	759
24.14.9	Inversion	760
24.15	Cepstrum	762
	References	769

25 Nonlinear Optimization	773
25.1 Gradient-Based Optimization	775
25.1.1 The Steepest Descent Technique	775
25.1.2 Newton's Minimization Technique	777
25.1.3 Quasi-Newton or Large Step Gradient Techniques	779
25.1.4 Conjugate Gradient Methods	793
25.2 Gradient-Free Optimization	803
25.2.1 Search Methods	804
25.2.2 Gradient-Free Conjugate Direction Methods	804
25.3 The Line Search Sub-Problem	809
25.4 Practical Considerations	810
25.4.1 Large-Scale Optimization	810
25.4.2 Numerical Stability	813
25.4.3 Nonsmooth Optimization	814
25.5 Constrained Optimization	814
25.5.1 The Lagrangian and Lagrange Multipliers	817
25.5.2 Duality	831
25.6 Global Convergence	835
References	836
26 Standards	841
26.1 Standard Audio Formats	842
26.1.1 Linear PCM (Uniform PCM)	842
26.1.2 μ -Law PCM (PCMU)	843
26.1.3 A-Law (PCMA)	843
26.1.4 MP3	843
26.1.5 HE-AAC	844
26.1.6 OGG Vorbis	844
26.1.7 ADPCM (G.726)	845
26.1.8 GSM	845
26.1.9 CELP	847
26.1.10 DTMF	848
26.1.11 Others Audio Formats	848
26.2 Standard Audio Encapsulation Formats	849
26.2.1 WAV	849
26.2.2 SPHERE	850
26.2.3 Standard Audio Format Encapsulation (SAFE)	850
26.3 APIs and Protocols	854
26.3.1 SVAPI	855
26.3.2 BioAPI	855
26.3.3 VoiceXML	856
26.3.4 MRCP	857
26.3.5 Real-time Transport Protocol (RTP)	858
26.3.6 Extensible MultiModal Annotation (EMMA)	858
References	859

Bibliography	861
Bibliography	861
Solutions	901
Index	909

Acronyms and Abbreviations

ADPCM	Adaptive Differential Pulse Code Modulation
AEP	Asymptotic Equipartition Property
AGN	Automatic Gain Normalization
AHC	Agglomerative Hierarchical Clustering
ANSI	American National Standards Institute
API	Application Programming Interface
ASR	Automatic Speech Recognition
BFGS	Broyden-Fletcher-Goldfarb-Shanno
BIC	Bayesian Information Criterion
BioAPI	Biometric Application Programming Interface
CBEFF	Common Biometric Exchange Formats Framework
CDMA	Code Division Multiple Access
CELP	Code Excited Linear Prediction
CHN	Cepstral Histogram Normalization
CMA	Constant Modulus Algorithm
CMN	Cepstral Mean Normalization
CMS	Cepstral Mean Subtraction
CMVN	Cepstral Mean and variance Normalization
CNG	Comfort Noise Generation
CoDec	Coder/Decoder
CS-ACELP	Conjugate Structure Algebraic Code Excited Linear Prediction
dB	deci Bel (decibel)
DC	Direct Current
DCF	Detection Cost Function
DCT	Discrete Cosine Transform
DET	Detection Error Trade-Off
DFP	Davidon-Fletcher-Powell
DHC	Divisive Hierarchical Clustering
DPCM	Differential Pulse Code Modulation
DTMF	Dual Tone Multi-Frequency

EER	Equal-Error Rate
e.g.	exempli gratia (for example)
EIH	Ensemble Interval Histogram
ELRA	European Language Resources Association
EM	Expectation Maximization
EMD	Empirical Mode Decomposition
EMMA	Extensible Multimodal Annotation
ETSI	European Telecommunications Standards Institute
FA	Factor Analysis
FAR	False Acceptance Rate
FBI	Federal Bureau of Investigation
FFT	Fast Fourier Transform
FRR	False Rejection Rate
FTP	File Transfer Protocol
GLR	General Likelihood Ratio
GMM	Gaussian Mixture Model(s)
GrXML	Grammar eXtensible Markup Language
GSM	Groupe Spécial Mobile <i>or</i> Global System for Mobile Communications
GSM-EFR	GSM Enhanced Full Rate
HE-AAC	High Efficiency Advanced Audio Coding
HEQ	Histogram Equalization
HME	Hierarchical Mixtures of Experts
HMM	Hidden Markov Model(s)
H-Norm	Handset Normalization
HTER	Half Total Error Rate
HTTP	HyperText Transfer Protocol
Hz	Hertz
IBM	International Business Machines
ID	Identity; Identification
iDEN	Integrated Digital Enhanced Network
i.e.	id est (that is)
IEC	International Electrotechnical Commission
IETF	Internet Engineering Task Force
IFG	Inferior Frontal Gyrus (of the Brain)
<i>i.i.d.</i>	Independent and Identically Distributed (Description of a type of Random Variable)
IMF	Intrinsic Mode Function
INCITS	InterNational Committee for Information Technology Standards
ISO	International Organization for Standardization
ISV	Independent Software Vendor
ITU	International Telecommunications Union
ITU-T	ITU Telecommunication Standardization Sector
JFA	Joint Factor Analysis
JTC	Joint ISO/IEC Technical Committee
IVR	Interactive Voice Response

KLT	Karhunen-Loève Transformation
LBG	Linde-Buzo-Gray
LFA	Latent Factor Analysis
kHz	kilo-Hertz
LDC	Linguistic Data Consortium
LAR	Log Area Ratio
LLN	Law of Large Numbers
LLR	Log-Likelihood Ratio
LPC	Linear Predictive Coding, also, Linear Predictive Coefficients
LPCM	Linear Pulse Code Modulation
MAP	Maximum A-Posteriori
MFCC	Mel Frequency Cepstral Coefficients
MFDWC	Mel Frequency Discrete Wavelet Coefficients
MIT-LL	Massachusetts Institute of Technology's Lincoln Laboratories
MLE	Maximum Likelihood Estimation or Maximum Likelihood Estimate
MLLR	Maximum Likelihood Linear Regression
MMIE	Maximum Mutual Information Estimation
MPEG	Moving Picture Experts Group
MRCP	Media Resource Control Protocol
NAP	Nuisance Attribute Projection
N.B.	Nota Bene (Note Well) – Note that
NIST	National Institute of Standards and Technology
NLSML	Natural Language Semantics Markup Language
NLU	Natural Language Understanding
OGI	Oregon Graduate Institute
PAM	Pulse Amplitude Modulation (Sampler)
PARCOR	Partial Correlation
PCA	Principal Component Analysis
PCM	Pulse Code Modulation
PCMA	A-Law Pulse Code Modulation
PCMU	μ -Law Pulse Code Modulation
PDC	Personal Digital Cellular
ppm	Parts per Million
pRAM	Probabilistic Random Access Memory
PSTN	Public Switched Telephone Network
PWM	Pulse Width Modulation (Sampler)
PWPAM	Pulse Width Pulse Amplitude Modulation (Sampler)
QCELP	Qualcomm Code Excited Linear Prediction
Q.E.D.	Quod Erat Demonstrandum (That which was to be Demonstrated)
QOS	Quality of Service
rad.	radians
RASTA	RelAtive SpecTrAl
RBF	Radial Basis Function
RFC	Request for Comments

RIFF	Resource Interchange File Format
RNN	Recurrent Neural Network
ROC	Receiver Operator Characteristic
RTP	Real-time Transport Protocol
SAFE	Standard Audio Format Encapsulation
SC	Subcommittee
SI	Système International
SIMM	Sequential Interacting Multiple Models
SIP	Session Initiation Protocol
SIV	Speaker Identification and Verification
LLN	Strong Law of Large Numbers
SPHERE	SPeECH HEader REsources
SPI	Service Provider Interface
SRAPI	Speech Recognition Application Programming Interface
SSML	Speech Synthetic Markup Language
SVAPI	Speaker Verification Application Programming Interface
SVM	Support Vector Machine(s)
TCP	Transmission Control Protocol
TD-SCDMA	Time Division Synchronous Code Division Multiple Access
TLS	Transport Layer Security
TDMA	Time Division Multiple Access
TDNN	Time-Delay Neural Network
T-Norm	Test Normalization
TTS	Text To Speech
U8	Unsigned 8-bit Storage
U16	Unsigned 16-bit Storage
U32	Unsigned 32-bit Storage
U64	Unsigned 64-bit Storage
UDP	User Datagram Protocol
VAD	Voice Activity Detection
VAR	Value Added Reseller
VB	Variational Bayesian Technique
VBWG	Voice Browser Working Group
VoiceXML	Voice eXtensible Markup Language
VoIP	Voice Over Internet Protocol
VQ	Vector Quantization
W3C	World Wide Web Consortium
WG	Workgroup
WCDMA	Wideband Code Division Multiple Access
WCDMA HSPA	Wideband Code Division Multiple Access High Speed Packet Access
WLLN	Weak Law of Large Numbers
XML	eXtensible Markup Language

Nomenclature

In this book, lower-case bold letters are used to denote vectors and upper-case bold letters are used for matrices. For set, measure, and probability theory, as much as possible, special style guidelines have been used such that the letter X when written as \mathcal{X} signifies a set and when written as \mathfrak{X} is a class of (sub)sets. The following is a list of symbols used in the text:

$\{\emptyset\}$	Empty Set
$\overline{(\alpha + i\beta)}$	Complex Conjugate of $(\alpha + i\beta)$ equal to $(\alpha - i\beta)$
$ \cdot $	Determinant of \cdot .
$(\mathbf{a})_{[i]}$	i^{th} element of vector \mathbf{a} .
$(\mathbf{A})_{[i][j]}$	Element in row i and column j of matrix \mathbf{A} .
$(\mathbf{A})_{[i]}$	Column i of matrix \mathbf{A} .
$*$	Convolution, e.g., $g * h$.
\circ	Correlation (Cross-Correlation), e.g., $g \circ h$, $g \circ g$.
$\hat{\cdot}$	Estimate of \cdot .
\wedge	Logical And
\vee	Logical Or
\mapsto	Maps to, e.g. $\mathcal{R}^N \mapsto \mathcal{R}^M$
\leftrightarrow	Mutual Mapping (used for signal/transform pairs, e.g. $h(t) \leftrightarrow H(s)$).
\therefore	Therefore
$\overset{\text{R}}{\equiv}$	Equivalent with respect to equivalence relation R.
\sim	Distributed According to \dots (a Distribution).
\preceq	$\mathbf{a} \preceq \mathbf{b}$ is read, \mathbf{a} precedes \mathbf{b} – i.e. in an ordered set of vectors.
\prec	$\mathbf{a} \prec \mathbf{b}$ is read, \mathbf{a} strictly precedes \mathbf{b} – i.e. in an ordered set of vectors.
\succeq	$\mathbf{a} \succeq \mathbf{b}$ is read, \mathbf{a} succeeds \mathbf{b} – i.e. in an ordered set of vectors.
\succ	$\mathbf{a} \succ \mathbf{b}$ is read, \mathbf{a} strictly succeeds \mathbf{b} – i.e. in an ordered set of vectors.
\bar{x}	Mean (Expected Value) of x
\mathcal{A}	A generic set.
\mathcal{A}^c	Complement of set \mathcal{A} .
$\mathcal{A} \setminus \mathcal{B}$	The difference between \mathcal{A} and \mathcal{B} .

A	Jacobian matrix of optimization constraints with respect to \mathbf{x}
\mathcal{B}	A generic set.
\mathcal{B}_c	Center Frequency of a Critical Band
\mathcal{B}_w	Bandwidth of a Critical Band
\mathbb{C}	Set of Complex Numbers
\mathcal{C}	Cost Function
\mathcal{C}^n	n -dimensional Complex Space
D	Dimension of the feature vector
Δ	Step Change
\mathcal{D}	Domain of a Function
$\mathbf{I}_{\mathcal{A}}(x)$	Characteristic function of $\mathcal{A} \in \mathcal{X}$ for random variable X
$\mathcal{D}_F(\cdot \leftrightarrow \cdot)$	f -Divergence
$\mathcal{D}_J(\cdot \leftrightarrow \cdot)$	Jeffreys Divergence
$\mathcal{D}_{KL}(\cdot \rightarrow \cdot)$	Kullback-Leibler Divergence
$d_E(\cdot, \cdot)$	Euclidean Distance
$d_{WE}(\cdot, \cdot)$	Weighted Euclidean Distance
$d_H(\cdot, \cdot)$	Hamming Distance
$d_{He}(\cdot, \cdot)$	Hellinger's Distance
$d_M(\cdot, \cdot)$	Mahalanobis Distance
$\nabla_{\mathbf{x}}E$	Gradient of E with respect to \mathbf{x}
$E(\cdot)$	Objective Function of Optimization
$\mathcal{E}\{\cdot\}$	Expectation of \cdot
e	Euler's Constant (2.7182818284...)
\mathbf{e}_n	Error vector
$\bar{\mathbf{e}}_N$	N -dimensional vector of all ones, i.e. $\bar{\mathbf{e}} : \mathcal{R}^1 \mapsto \mathcal{R}^N$ such that, $(\bar{\mathbf{e}}_N)_{[n]} = 1$ for all $n = \{1, 2, \dots, N\}$
$\hat{\mathbf{e}}_k$	Unit vector whose k^{th} element is 1 and all other elements are 0
$\exp\{\cdot\}$	Exponential function ($e^{\{\cdot\}}$)
ϕ	Sample Space of the Parameter Vector, $\boldsymbol{\varphi}$
$\boldsymbol{\varphi}_\gamma$	Parameter Vector for the cluster γ
$\boldsymbol{\Phi}$	Matrix of parameter vectors
F_s	Spectral Flatness
$\mathcal{F}\{\cdot\}$	Fourier Transform of \cdot
$\mathcal{F}^{-1}\{\cdot\}$	Inverse Fourier Transform of \cdot
\mathfrak{F}	A Field
$\mathcal{I}_F(\boldsymbol{\varphi} \mathbf{x})$	Fisher Information matrix for parameter vector $\boldsymbol{\varphi}$ given \mathbf{x}
f	Frequency measured in Hertz ($\frac{\text{cycles}}{s}$)
f_c	Nyquist Critical Frequency measured in Hertz ($\frac{\text{cycles}}{s}$)
f_s	Sampling Frequency measured in Hertz ($\frac{\text{cycles}}{s}$)
Γ	Number of clusters – mostly Gaussian clusters
γ	Cluster index – mostly for Gaussian clusters
$\boldsymbol{\gamma}_{n_c}$	Column n_c of Jacobian matrix (\mathbf{J}) of optimization constraints
G	Hessian Matrix

\mathbf{g}	Gradient Vector
$\mathcal{H}(p)$	Entropy
$\mathcal{H}(p q)$	Conditional Entropy
$\mathcal{H}(p, q)$	Joint Entropy
$\mathcal{H}(p \rightarrow q)$	Cross Entropy
\mathbf{H}	Inverse Hessian Matrix
\mathcal{H}	Hilbert Space
\mathfrak{H}	Borel Field of the Borel Sets in Hilbert Space
\mathcal{H}_p	Pre-Hilbert Space
\mathfrak{H}_p	Borel Field of the Borel Sets in Pre-Hilbert Space
H_0	Null Hypothesis
H_1	Alternative Hypothesis
$H(f)$	Fourier Transform of the signal $h(t)$
$H(s)$	Laplace Transform of the signal $h(t)$
$H(s)$	Any Generic Function of a Complex Variable
$H(\omega)$	Fourier Transform of the signal $h(t)$ in Terms of the Angular Frequency ω
H_{kl}	Discrete Fourier Transform of the sampled signal h_{nl} in frame l for the linear frequency index k
\check{H}_{ml}	Mel-scale Discrete Fourier Transform of the sampled signal h_{nl} in frame l for the Mel frequency index m
$h(t)$	A Continuous Function of Time or a Continuous Signal
$\tilde{h}(p)$	Differential Entropy (Continuous Entropy)
$\tilde{h}(p \rightarrow q)$	Differential Cross Entropy (Continuous Cross Entropy)
I_0	Standard Intensity Threshold for Hearing
I	Intensity of Sound
I_r	Relative Intensity of Sound
\mathcal{I}	Information
$\mathcal{I}(X; Y)$	Mutual Information between Random Variables X and Y
$\mathcal{I}_J(X; Y)$	Jeffrey's Mutual Information between Random Variables X and Y
\mathbb{I}	Set of Imaginary Numbers
\mathbf{I}	Identity Matrix
$\mathcal{I}m$	The Imaginary part of variable $\{s : s \in \mathbb{C}\}$
\mathbf{I}_N	N -dimensional Identity Matrix
i	The Imaginary Number ($\sqrt{-1}$)
<i>iff</i>	If and Only If (\iff)
<i>inf</i>	Infimum
$\mathcal{K}(t, s)$	Kernel Function of t and s used in Integral Transforms
$\mathbf{\Lambda}$	Diagonal matrix of Eigenvalues
λ	Lebesgue Measure
$\tilde{\lambda}$	Wavelength
$\bar{\lambda}$	Forgetting Factor

$\hat{\lambda}$	Eigenvalue
λ	Lagrange Multiplier
L	Total number of frames
$\mathcal{L}(\boldsymbol{\varphi} \mathbf{x})$	Likelihood of $\boldsymbol{\varphi}$ given \mathbf{x}
$\mathcal{L}\{\cdot\}$	Laplace Transform of \cdot
$\mathcal{L}^{-1}\{\cdot\}$	Inverse Laplace Transform of \cdot
\mathcal{L}_p	Class of extended real valued p -integrable functions
l	Frame Index
$\ell(\boldsymbol{\varphi} \mathbf{x})$	Log-Likelihood of $\boldsymbol{\varphi}$ given \mathbf{x}
$\ln(\cdot)$	Napierian Logarithm, Natural Logarithm, or Hyperbolic Logarithm ($\log_e(\cdot)$)
$\log(\cdot)$	Common Logarithm ($\log_{10}(\cdot)$)
$\boldsymbol{\mu}$	Mean Vector
$\hat{\boldsymbol{\mu}}$	Sample mean vector, as a shortcut for $\bar{X} _N$
$\hat{\boldsymbol{\mu}}_\gamma$	Sample mean vector for cluster γ
M	Number of Models, number of critical bands
M	Number of samples in a partition of the Welch PSD computation
M	Dimension of the parameter vector
\mathcal{M}	Matrix of the weights for mapping the linear frequency to the Mel scale critical filter bank frequencies
$\mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$	Gaussian or Normal Distribution with mean $\boldsymbol{\mu}$ and Variance-Covariance $\boldsymbol{\Sigma}$
N	Window size
N	Number of samples
N	Number of hypotheses
n	Sample index which is not necessarily time aligned – see t for time aligned sample index
N_γ	Number of samples associated with cluster γ
N_s	Number of samples associated with state s
\mathbb{N}	The set of Natural Numbers
O	Observation random variable
\mathcal{O}	Observation sample space
\mathcal{O}	Bachmann-Landau asymptotic notation – Big-O notation
\mathcal{D}	Borel Fields of the Borel Sets of sample space \mathcal{O}
o	An observation sample
ϖ	Pulsewidth of Pulse Amplitude Modulation Sampler
$\varpi(o s)$	Penalty (loss) associated with decision o conditioned on state s
$\varpi(o x)$	Conditional Risk in Bayesian Decision Theory
f_0	Pitch
$\boldsymbol{\Pi}$	Penalty matrix in Bayesian Decision Theory.
P	Probability
P	Pressure Differential
P_0	Pressure Threshold
\mathcal{P}	Total Power

\mathcal{P}_d	Power Spectral Density
\mathcal{P}_d°	Power Spectral Density in Angular Frequency
p	Probability Distribution
\mathbf{p}	Training patten index for a Neural Network
q	Probability Distribution
\mathbb{R}	Set of Real Numbers
R	Redundancy
$\mathcal{R}(h)$	Range of Function h – Set of values which function h may take on
$\Re(s)$	The Real part of variable $\{s : s \in \mathbb{C}\}$
\mathcal{R}^n	n -dimensional Euclidean Space
Σ	Covariance (Variance-Covariance) Matrix
$\hat{\Sigma}$	Biased Sample Covariance (Variance-Covariance) Matrix
$\tilde{\Sigma}$	Unbiased Sample Covariance (Variance-Covariance) Matrix
$\hat{\Sigma}_\gamma$	Biased Sample Covariance Matrix for cluster γ
\textcircled{S}	Number of States
S	State Random variable
\mathcal{S}	State sample space
\mathcal{G}	State Borel Field of the Borel Sets of sample space \mathcal{S}
$\mathbf{S} _N$	Second Order Sum ($\sum_{i=1}^N \mathbf{x}_i \mathbf{x}_i^T$)
s	A sample of the state random variable
$\mathbf{s} _N$	First Order Sum ($\sum_{i=1}^N \mathbf{x}_i$)
sup	Supremum
$\zeta(\boldsymbol{\phi} \mathbf{x})$	Score Statistic (Fisher Score) for parameters vector $\boldsymbol{\phi}$ given \mathbf{x}
T	Total Number of Samples, and sometimes the Sampling Period
t	Sample index in time
T_c	Nyquist Critical Sampling Period
T_s	Sampling Period
$\hat{\mathbf{u}}$	Unit Vector
ω	Angular Frequency measured in $\frac{\text{rad.}}{s}$
ω_c	Nyquist Critical Angular Frequency measured in $\frac{\text{rad.}}{s}$
ω_s	Angular Sampling Frequency measured in $\frac{\text{rad.}}{s}$
W_N	The Twiddle Factor used for expressing DFT ($e^{i\frac{2\pi}{N}}$)
W_N^{kn}	$W_N^{(k \times n)}$
\mathcal{E}	Seconds of shift in feature computation
\mathfrak{X}	Borel Field (the smallest σ -field) of the Borel Sets of Sample Space, \mathcal{X}
\mathcal{X}	Sample Space
\mathbf{x}	Feature Vector
$\mathcal{Z}\{\cdot\}$	z Transform of \cdot
$\mathcal{Z}^{-1}\{\cdot\}$	Inverse z Transform of \cdot
\mathbb{Z}	The Set of Integers
\mathbf{z}_k	Direction of the Inverse Hessian Update in Optimization

List of Figures

1.1	Open-Set Segmentation Results for a Conference Call <i>Courtesy of Recognition Technologies, Inc.</i>	9
1.2	Diagram of a Full Speaker Diarization System including Transcription for the Generation of an Indexing Database to be Used for Text+ID searches	20
1.3	Proctorless Oral Language Proficiency Testing <i>Courtesy of Recognition Technologies, Inc.</i>	22
1.4	Indexing Based on Audio	36
1.5	Indexing Based on Video	36
1.6	Speech Generation Model after Rabiner [55]	37
2.1	Sagittal section of Nose, Mouth, Pharynx, and Larynx; <i>Source:</i> <i>Gray's Anatomy [13]</i>	45
2.2	Sagittal Section of Larynx and Upper Part of Trachea; <i>Source:</i> <i>Gray's Anatomy [13]</i>	46
2.3	Coronal Section of Larynx and Upper Part of Trachea; <i>Source:</i> <i>Gray's Anatomy [13]</i>	46
2.4	Laryngoscopic View of the interior Larynx; <i>Source: Gray's</i> <i>Anatomy [13]</i>	46
2.5	The Entrance to the Larynx, Viewed from Behind; <i>Source: Gray's</i> <i>Anatomy [13]</i>	46
2.6	The External Ear and the Middle Ear; <i>Source: Gray's Anatomy [13]</i> .	49
2.7	The Middle Ear; <i>Source: Gray's Anatomy [13]</i>	49
2.8	The Inner Ear; <i>Source: Gray's Anatomy [13]</i>	49
2.9	A Typical Neuron	52
2.10	Sagittal Section of the Human Brain (<i>Source: Gray's Anatomy [13]</i>) .	55
2.11	MRI of the Left Hemisphere of the Brain	56
2.12	Left Cerebral Cortex (Inflated)	57
2.13	Left Cerebral Cortex (Flattened)	57

2.14	Left Hemisphere of the Human Brain (<i>Modified from: Gray's Anatomy [13]</i>)	57
2.15	Centers of the Lateral Brodmann Areas	58
2.16	Areas of Speech Production in the Human Brain	60
2.17	Areas of Speech Understanding in the Human Brain	61
2.18	Speech Generation and Perception – Adapted From Figure 1.6	65
2.19	Language Production and Understanding Regions in the Brain (<i>Basic Figure was adopted from Gray's Anatomy [13]</i>)	66
2.20	Auditory Mapping of the Brain and the Cochlea (<i>Basic figures were adopted from Gray's Anatomy [13]</i>)	67
2.21	The Auditory Neural Pathway – Relay Path toward the Auditory Cortex	68
2.22	Speech Signal Transmission between the Ears and the Auditory Cortex – See Figure 2.21 for the connection into the lower portion indicated by clouds	69
2.23	The connectivity and relation among the audio cortices and audio perception areas in the two hemispheres of the cerebral cortex	70
2.24	Corpus Callosum, which is in charge of communication between the two hemispheres of the brain	71
3.1	Sampling of a Simple Sine Signal at Different Sampling Rates; $f =$ Signal Frequency $f_s =$ Sampling Frequency – The Sampling Rate starts at $2f$ and goes up to $10f$	78
3.2	<i>sinc</i> function which is known as the cardinal function of the signal – f_c is the Nyquist Critical Frequency and ω_c is the corresponding Nyquist Angular Frequency ($\omega_c = 2\pi f_c$)	83
3.3	Portion of a speech waveform sampled at $f_s = 22050$ Hz – Solid line shows the signal quantized into 11 levels and the dots show original signal	86
3.4	Speech Waveform sampled at $f_s = 22050$ Hz	87
3.5	Narrowband spectrogram using ~ 23 ms widows (43Hz Band)	88
3.6	Wideband spectrogram using ~ 6 ms widows (172Hz Band)	88
3.7	Z-IH-R-OW	89
3.8	W-AH-N	89
3.9	T-UW	89
3.10	TH-R-IY	89
3.11	F-OW-R	90
3.12	F-AY-V	90
3.13	S-IH-K-S	90
3.14	S-EH-V-AX-N	90
3.15	EY-T	90
3.16	N-AY-N	90
3.17	Formants shown for an elongated utterance of the word [try] – see Figure 4.29 for an explanation.	91
3.18	Adult male (44 years old)	91

3.19	Male child (2 years old)	91
3.20	Uniform Rate Pulse Amplitude Modulation Sampler. top: Waveform plot of a section of a speech signal. middle: Pulse Train $p(t)$ at $T_s = 5 \times 10^{-4}$ s (2kHz) and $\varpi = \frac{T_s}{10}$ bottom: Pulse Amplitude Modulated samples overlaid with the original signal for reference.	93
3.21	Pulse Width Modulation Sampler. top: Waveform plot of a section of a speech signal. bottom: Pulse Width Modulated samples overlaid with the original signal for reference.	94
3.22	Pulse Amplitude Modulation Sampler Block Diagram (after [10])	94
3.23	Magnitude of the <i>complex Fourier series coefficients</i> of a uniform-rate fixed pulsewidth sampler	97
3.24	Reflections in the Laplace plane due to folding of the Laplace Transform of the output of an ideal sampler – x marks a set of poles which are also folded to the higher frequencies	100
3.25	The first $\frac{1}{2}$ second of the signal in Figure 3.28	101
3.26	Original signal was subsampled by a factor of 4 with no filtering done on the signal	101
3.27	The original signal was subsampled by a factor of 4 after being passed through a low-pass filter	101
3.28	“Sampling Effects on Fricatives in Speech” (Sampling Rate: 22 kHz)	104
3.29	“Sampling Effects on Fricatives in Speech” (Sampling Rate: 8 kHz)	104
4.1	Fundamental Frequencies for Men, Women and Children while uttering 10 common vowels in the English Language – Data From [18]	113
4.2	Formant 1 Frequencies for Men, Women and Children while uttering 10 common vowels in the English Language – Data From [18]	113
4.3	Formant 2 Frequencies for Men, Women and Children while uttering 10 common vowels in the English Language – Data From [18]	113
4.4	Formant 3 Frequencies for Men, Women and Children while uttering 10 common vowels in the English Language – Data From [18]	113
4.5	Position of the 10 most common vowels in the English Language as a function of formants 1 and 2 – Average Male Speaker	114
4.6	Position of the 10 most common vowels in the English Language as a function of formants 1 and 2 – Average Female Speaker	114
4.7	Position of the 10 most common vowels in the English Language as a function of formants 1 and 2 – Average Child Speaker	114
4.8	Position of the 10 most common vowels in the English Language as a function of formants 1 and 2 – Male, Female and Child	114
4.9	Persian ingressive nasal velaric fricative (click), used for negation – colloquial “No”	123

4.10	bead /bi:d/	
	(In an American Dialect of English)	125
4.11	bid /bɪd/	
	(In an American Dialect of English)	125
4.12	bayed /beɪd/	
	(In an American Dialect of English)	125
4.13	bed /bɛd/	
	(In an American Dialect of English)	125
4.14	bad /bæd/	
	(In an American Dialect of English)	126
4.15	body /bɑ:di/	
	(In an American Dialect of English)	126
4.16	bawd /bə:d/	
	(In an American Dialect of English)	126
4.17	Buddhist /bʊ dist/	
	(In an American Dialect of English)	126
4.18	bode /boʊ d/	
	(In an American Dialect of English)	127
4.19	booed /bu:d/	
	(In an American Dialect of English)	127
4.20	bud /bʌd/	
	(In an American Dialect of English)	127
4.21	bird /bɜ:d/	
	(In an American Dialect of English)	127
4.22	bide /baɪd/	
	(In an American Dialect of English)	128
4.23	bowed /baʊ d/	
	(In an American Dialect of English)	128
4.24	boyd /bɔ:d/	
	(In an American Dialect of English)	128
4.25	Vowel Trapezoid for the Persian Language	130
4.26	[try] Decisive Imperative – Short and powerful	134
4.27	[try] Imperative with a slight interrogative quality – short and an imperative; starts in the imperative tone and follows with an interrogative ending	134
4.28	[try] Imperative but with a stronger interrogative quality – longer and the pitch level rises, it is sustained and then it drops	134
4.29	Imperative in a grammatical sense, but certainly interrogative in tone – much longer; the emphasis is on the sustained diphthong at the end with pitch variation by rising, an alternating variation and a final drop	134
4.30	Mandarin word, Ma (Mother)	136
4.31	Mandarin word, Ma (Hemp)	136
4.32	Mandarin Word, Ma (Horse)	136
4.33	Mandarin Word, Ma (Scold)	136

4.34	construct of a typical syllable, [tip]	139
5.1	Pitch versus Frequency for frequencies of up to 1000 Hz	147
5.2	Pitch versus Frequency for the entire audible range	147
5.3	Block Diagram of a typical Sampling Process for Speech – Best Alternative	152
5.4	Block Diagram of a typical Sampling Process for Speech – Alternative Option	152
5.5	Block Diagram of a typical Sampling Process for Speech – Alternative Option	152
5.6	The power spectral density of the original speech signal sampled at 44100 Hz using the Welch PSD estimation method	154
5.7	The power spectral density of the pre-emphasized speech signal sampled at 44100 Hz using the Welch PSD estimation method	154
5.8	The spectrogram of the original speech signal sampled at 44100 Hz	154
5.9	The spectrogram of the pre-emphasized speech signal sampled at 44100 Hz	154
5.10	Block diagram of the human speech production system viewed as a control system	158
5.11	Frame of audio $N = 256$	161
5.12	Hi-Pass filtered Frame $N = 256$	161
5.13	Pre-Emphasized Frame of audio $N = 256$	162
5.14	Windowed Frame $N = 256$	162
5.15	Hamming Window	164
5.16	Hann Window	164
5.17	Welch Window	164
5.18	Triangular Window and its spectrum	166
5.19	Blackman Window ($\alpha = 0.16$) and its spectrum	166
5.20	Gauss Window ($\sigma = 0.4$) and its spectrum	166
5.21	The first 24 critical frequencies given by the Scale on the horizontal axis, On the left vertical axis, the corresponding Mels and on the right, the corresponding $\log(\text{frequency in Hz})$	168
5.22	Shape of the Mel filter bank weights for a 24-filter system with an 8000 Hz sampling frequency. Not all the frequency values for the lower frequency centers have been written since the numbers would have merged and would not be legible.	171
5.23	Power Spectrum of the Frame $N = 256$	171
5.24	Power Spectrum in the Mel Frequency Domain $N = 256, M = 24$	171
5.25	First 21 Mel Frequency Cepstral Coefficients	172
5.26	Mean value of the MFCC vectors over 260 frames	174
5.27	Mean-subtracted MFCC vector	174
5.28	Spectrogram of the audio being analyzed for MFCC computation	175
5.29	Trajectory of the first two MFCC components over the whole utterance	175

5.30	Shortpass lifted Trajectory of the first two MFCC components over the whole utterance	175
5.31	Cylinder Model of the Vocal Tract; <i>Source: Flanagan [21]</i>	186
5.32	Concentric cylinder model of the vocal tract	188
5.33	Physical interpretation of the reflection coefficients	188
5.34	Perceptual Linear Predictive Analysis due to Hermansky [30]	190
6.1	$\mathcal{A} \cap \mathcal{B} = \{\emptyset\}$ (Disjoint) (Mutually Exclusive)	206
6.2	Intersection, $\mathcal{A} \cap \mathcal{B}$ $(\mathcal{A} \wedge \mathcal{B})(\mathcal{A}, \mathcal{B})$	206
6.3	Union, $\mathcal{A} \cup \mathcal{B}$ $(\mathcal{A} \vee \mathcal{B})(\mathcal{A} + \mathcal{B})$	206
6.4	Complement of \mathcal{A} , $\mathcal{A}^c = \mathcal{X} \setminus \mathcal{A}$ $(!\mathcal{A})(\mathcal{A}')$	206
7.1	Information Flow in a Single Direction of a Generic Communication System	265
7.2	Decomposition of three choices into a binary set of choices – after Shannon [18]	271
7.3	Entropy of a Bernoulli Process	273
7.4	A Memoryless (Zero-Memory) Channel	283
9.1	Hypothesis Testing Logic for Speaker Verification	316
9.2	Boundaries of three discriminant functions based on normal density likelihoods and equal a-priori probability classes	325
9.3	Path of a sample complex binary question – it always has only two final outcomes.	336
13.1	Alternative Evolutionary Paths of the Hidden Markov Model	412
13.2	Two Coins, Unifilar Single Memory Markov Model	417
13.3	Two Coins, Non-Unifilar Single Memory Markov Model (HMM)	419
13.4	Basic HMM Element	420
13.5	Simplification using a Null Transition	422
13.6	HMM of an Average Phone	423
13.7	HMM of Example 13.4	425
13.8	Possible Paths for generating $y_1 = a$ in Example 13.4	425
13.9	Possible Paths for generating $\{y\}_1^2 = aa$ in Example 13.4	426
13.10	Merged paths for generating $\{y\}_1^2 = aa$ in Example 13.4	427
13.11	Trellis diagram for the output sequence, $\{y\}_1^4 = \{aabb\}$ being generated by the HMM of Figure 13.7	428
13.12	α computation for the HMM of Example 13.4	429
13.13	Viterbi maximum probability path computation for the HMM of Example 13.4	432
13.14	HMM of Example 13.5 with maximum entropy initial distributions	437
13.15	Trellis of Example 13.5 with maximum entropy initial distributions	438

13.16	HMM of Example 13.5 with recomputed distributions and transition probabilities after one iteration of Baum-Welch	438
13.17	Trellis of Example 13.5 with recomputed a -posteriori output-transition probabilities	439
13.18	Convergence of the likelihood of the HMM given the sequence, $\{abaa\}$, to a local maximum, as related to Example 13.5	440
13.19	Configuration of the locally converged HMM model for sequence $\{abaa\}$ in Example 13.5	440
13.20	A slight deviation from maximum entropy in the initial distribution of the HMM of Example 13.5	441
13.21	Configuration of the globally converged HMM model for sequence $\{abaa\}$ in Example 13.5	441
13.22	Convergence of the likelihood of the HMM given the sequence, $\{abaa\}$, for two different initial conditions: 1. maximum entropy and 2. slight perturbation from maximum entropy	442
13.23	Plot of the polynomial of Equation 13.102	457
13.24	State diagram of the smoothing model of Equation 13.99	458
13.25	Trellis diagram of the smoothing model of Equation 13.99	458
13.26	Repetitive trellis component ending in a confluent state	459
13.27	Convergence of the forward-backward estimate of α to the value found earlier in Figure 13.23	460
13.28	State diagram of the mixture coefficients for more than two distributions	461
14.1	A Perceptron	466
14.2	Generic L-Layer Feedforward Neural Network	467
14.3	Generic Recurrent Neuron	476
14.4	Generic Time-Delay Neural Network	478
14.5	Feedforward neural network architecture used for the <i>exclusive OR</i> (<i>XOR</i>) logic	481
15.1	Linearly Separable 2-Class Problem	490
18.1	Spectrogram of audio used in the filtering and liftering exercises	566
18.2	Top: MFCC components (c_1 and c_2) throughout the utterance; Bottom: Power Spectral Density of c_1 and c_2	568
18.3	Top: c_1 and c_2 after CMN; Bottom: Power Spectral Density of c_1 and c_2 after CMN	568
18.4	Top: c_1 and c_2 after CMVN; Bottom: Power Spectral Density of c_1 and c_2 after CMVN	569
18.5	Impulse response of the ARMA lifter	570
18.6	Impulse response of the RASTA lifter	570
18.7	Top: c_1 and c_2 after RASTA Liftering; Bottom: Power Spectral Density of c_1 after RASTA Liftering	572

18.8	Top: c_1 and c_2 after CMN and ARMA Liftering; Bottom: Power Spectral Density of c_1 after CMN and ARMA Liftering	572
18.9	Adult male (44 years old)	573
18.10	Male child (2 years old)	573
18.11	Impulse response of the shortpass lifter	574
18.12	c_1 and c_2 after applying CMN followed by a shortpass lifter	574
19.1	Sample Receiver Operating Characteristic (ROC) curve representing the same data which was used to plot the DET Curve of Figure 19.2	591
19.2	Sample Detection Error Tradeoff (DET) curve representing the same data which was used to generate the ROC Curve of Figure 19.1	592
19.3	Sample Histogram of the rank of the target speaker from a 78-speaker database with 78 tests using the same data as was used for the verification task in Figures 19.1 and 19.2	594
20.1	Number of Days Between Test 1 and 2	596
20.2	Number of Days Between Test 2 and 3	596
20.3	Identification Time Lapse – Usual Enrollment	599
20.4	Verification Time Lapse using Usual Enrollment	599
21.1	Identification Time Lapse – Augmented-Data Enrollment	602
21.2	Verification using Augmented-Data Enrollment	602
21.3	Identification Time Lapse	603
21.4	Verification using 5 iteration Adaptation	604
21.5	Verification using 1 iteration Adaptation	605
22.1	The Qualitative Relation between Linguistic Knowledge used in a Speech Recognition System and its Performance.	621
23.1	Polyhedral Cone Spanned by the columns of \mathbf{A} .	645
24.1	Representation of a Number s_0 in the Complex Plane	648
24.2	Representation of a Circle in \mathbb{C}	648
24.3	Point of ordinary discontinuity at $t = t_0 = 2$	654
24.4	Point of ordinary discontinuity at $t = t_0 = 0$ ($h(t) = \frac{\sin(t)}{ t }$)	654
24.5	Point of Oscillatory Discontinuity at $t = t_0 = 0$ ($h(t) = \frac{1}{t-0}$)	656
24.6	More Detailed Viewpoint of the Oscillatory Discontinuity at $t = t_0 = 0$	656
24.7	Graphic representation of the inequality in Equation 24.46. $h(t)$ is the convex function of interest and $f(t)$ describes the line in this figure.	659
24.8	Graphic representation of the inequality in Equation 24.63. $h(t)$ is the concave function of interest and $f(t)$ describes the line in this figure.	662
24.9	Integration Path of Multiply Connected Contours	677

24.10 Individual Contour Paths Used by the Cauchy Integral Theorem	677
24.11 Contour of Integration for Cauchy Integral Formula	678
24.12 Taylor Series Convergence for an Analytic Function	683
24.13 Laurent Series Annular Region of Convergence for an Analytic Function	683
24.14 The Haar Scale Function $\varphi(t)$	717
24.15 The Haar Mother Wavelet $\psi(t)$	717
24.16 Convergence Region for the Laplace Transform	720
24.17 Original signal overlaid with a square window of $60ms$ width at $t = 80ms$	740
24.18 Windowed Signal	740
24.19 Original signal overlaid with a square window of $60ms$ width at $t = 80ms$ with a normalized window using Equation 24.474	742
24.20 Windowed signal with a normalized window using Equation 24.474	742
24.21 The corresponding extrema in the Laplace and z planes	754
24.22 Waveform and spectrogram of the signal and its echo at a delay of 0.3 s and a reflection factor of 0.4 , given by Equations 24.602 and 24.603	765
24.23 A small portion of the signal in the vicinity of the arrival of the echo	765
24.24 Spectrum of the signal, showing the 4 periodic components	765
24.25 Cepstrum of the signal, showing a peak at the moment of arrival of the echo	765
25.1 Stationary Points of a Function of Two Variables	774
25.2 Contour plot of the function of Figure 25.1	776
25.3 Flowchart of Davidon's Quasi-Newton Method – Part 1	794
25.4 Flowchart of Davidon's Quasi-Newton Method – Part 2	795
25.5 Flowchart of Davidon's Quasi-Newton Method – Part 3	796
25.6 Powell's Convergence to a Set of Conjugate Directions	808

List of Tables

1.1	DNA Nucleotides	25
1.2	Sample Audio/Video and Fusion Results For Multimodal Speaker Recognition [73]	36
4.1	Types of Clicks in Nama (South African Language) Data was extracted from samples in the CD-ROM accompanying [8]	121
4.2	Examples of whispered alterations in Comanche – the circle under a vocoid changes it to a whispered phonation	126
4.3	Examples, in English, where stress changes the meaning of a word ..	137
4.4	Parallels between Pitch and Loudness Concepts	137
5.1	Frequency versus Pitch	148
14.1	XOR input/output for Problem 14.1	482
16.1	<i>The Hypothetical Menagerie</i> – an animal analogy of the behavior of speakers, defined by [21]	537
16.2	<i>The Biometric Menagerie</i> – an animal analogy of the behavior of speakers, defined by [54, 55]	538
24.1	Terminology analogs of the spectral domain, used in the cepstral domain	763
26.1	Mobile technology subscribers, worldwide, in the second quarter of 2009 according to GSMA [8]	846
26.2	Audio Interchange Scenarios	851
26.3	Macros	851
26.4	Audio Format Header	853

List of Definitions

3.1	Definition (Signal)	75
3.2	Definition (Signal)	75
3.3	Definition (Time-Dependent Signal)	75
3.4	Definition (Stationary and Non-Stationary Signals)	76
4.1	Definition (Phonetics)	107
4.2	Definition (phone)	108
4.3	Definition (Phoneme)	108
4.4	Definition (Allophone)	108
4.5	Definition (Phonology)	122
4.6	Definition (Prosodic Features)	132
4.7	Definition (Pitch)	132
4.8	Definition (Loudness (Sonority))	137
4.9	Definition (Co-Articulation)	140
5.1	Definition (Melody (Mel))	147
5.2	Definition (Bark)	148
5.3	Definition (Phon)	151
5.4	Definition (Sone)	151
5.5	Definition (Spectral Flatness)	169
6.1	Definition (Sample Space)	205
6.2	Definition (Event)	205
6.3	Definition (Class)	207
6.4	Definition (Countable Base)	207
6.5	Definition (Countable Space)	207
6.6	Definition (Countably Infinite Space)	207
6.7	Definition (Closure under a set operation)	208
6.8	Definition (Equivalence Relation)	208
6.9	Definition (Equivalence Class)	208
6.10	Definition (Quotient Set)	209

6.11 Definition (Partition)	209
6.12 Definition (Field, Algebra, or Boolean Algebra)	212
6.13 Definition (Field)	213
6.14 Definition (σ -Field (σ -algebra))	214
6.15 Definition (Borel Field)	214
6.16 Definition (Borel Sets)	214
6.17 Definition (Measurable Space)	214
6.18 Definition (Measure)	215
6.19 Definition (Lebesgue Measure)	215
6.20 Definition (Lebesgue Measurable Subsets)	215
6.21 Definition (Lebesgue Field)	215
6.22 Definition (Lebesgue Measure Space)	216
6.23 Definition (Complete Space)	216
6.24 Definition (Cartesian Product)	216
6.25 Definition (Metric Space)	217
6.26 Definition (Distance Between Subsets)	217
6.27 Definition (Convex Metric Space)	218
6.28 Definition (A Convex subset of the Euclidean Space)	218
6.29 Definition (Complete Metric Space)	218
6.30 Definition (Banach Space)	218
6.31 Definition (Inner Product Space (Dot Product Space))	219
6.32 Definition (Pre-Hilbert Space)	220
6.33 Definition (Hilbert Space)	220
6.34 Definition (Hilbert Metric Space)	221
6.35 Definition (Probability Measure)	222
6.36 Definition (Conditional Probability)	224
6.37 Definition (Statistical Independence)	226
6.38 Definition (Mutual Statistical Independence)	226
6.39 Definition (Continuous Random Variable)	227
6.40 Definition (Extended Real Valued Function)	228
6.41 Definition (Absolute Continuity of Two Measures)	229
6.42 Definition (Equivalence)	229
6.43 Definition (Almost Everywhere (Modulo))	229
6.44 Definition (Probability Density Function)	230
6.45 Definition (Probability Density Function)	231
6.46 Definition (Joint Probability Density Function)	233
6.47 Definition (Marginal Probability Density Function)	233
6.48 Definition (Cumulative Distribution Function)	235
6.49 Definition (\mathcal{L}_p Class of p -integrable Functions (Lebesgue space))	236
6.50 Definition (Schwarz's Inequality)	237
6.51 Definition (Transformation)	238
6.52 Definition (One-to-One Transformation)	238
6.53 Definition (Product of Transformation)	238
6.54 Definition (Inverse Image of a Transformation)	238
6.55 Definition (Measurable Transformation)	238

6.56	Definition (Expected Value (Expectation))	239
6.57	Definition (Expected Value or Mean)	240
6.58	Definition (Conditional Expectation)	240
6.59	Definition (Covariance)	245
6.60	Definition (Correlation Coefficient)	245
6.61	Definition (Excess Kurtosis)	246
6.62	Definition (Discrete Random Variable)	247
6.63	Definition (Bernoulli Random Variable)	247
6.64	Definition (Probability Distribution (Probability Mass Function))	247
6.65	Definition (Cumulative Probability Distribution)	247
6.66	Definition (Expected Value (Mean) of a Discrete Random Variable)	248
6.67	Definition (Expected Value of a Function of a Discrete Random Variable)	248
6.68	Definition (Variance of a Discrete Random Variable)	249
6.69	Definition (Weak Convergence (Convergence in Probability))	250
6.70	Definition (Strong Convergence (Almost Sure Convergence or Convergence with Probability 1))	251
6.71	Definition (Fundamental Sequence or a Cauchy Sequence)	251
6.72	Definition (Statistic)	251
6.73	Definition (Sufficient Statistic)	252
6.74	Definition (Efficiency of a Statistic)	252
6.75	Definition (Statistical Efficiency Criterion)	252
6.76	Definition (Efficient Statistic)	252
6.77	Definition (Consistent Statistic)	253
6.78	Definition (Arithmetic Mean)	257
6.79	Definition (Geometric Mean)	257
6.80	Definition (Harmonic Mean)	257
6.81	Definition (Quadratic Mean (Root Mean Square – RMS))	257
6.82	Definition (Sample Variance (Biased Estimator))	258
6.83	Definition (Sample Variance (Unbiased Estimator))	258
7.1	Definition (Discrete Source)	267
7.2	Definition (Discrete Memoryless Source)	267
7.3	Definition (Discrete Markov (Markoff) Source)	268
7.4	Definition (Unifilar Markov Source)	268
7.5	Definition (Non-Unifilar Markov Source)	268
7.6	Definition (Ergodic Sources)	268
7.7	Definition (Continuous Source)	269
7.8	Definition (Entropy)	272
7.9	Definition (Relative Entropy – as defined by Shannon)	273
7.10	Definition (Redundancy)	274
7.11	Definition (Joint Entropy)	275
7.12	Definition (Conditional Entropy)	277
7.13	Definition (Discrete Channel)	283
7.14	Definition (Discrete Memoryless Channel)	283

7.15 Definition (Binary Symmetric Channel) 284

7.16 Definition (Differential Cross Entropy) 290

7.17 Definition (Cross Entropy) 290

8.1 Definition (Euclidean Distance) 302

8.2 Definition (L_p Distance) 303

8.3 Definition (Weighted Euclidean Distance) 303

8.4 Definition (Mahalanobis Distance) 303

8.5 Definition (Bhattacharyya Divergence) 307

8.6 Definition (f -Divergence) 308

8.7 Definition (δ -Divergence) 309

9.1 Definition (Null Hypothesis) 313

9.2 Definition (Alternative Hypothesis) 313

9.3 Definition (Target Speaker (Reference Model)) 314

9.4 Definition (Test Speaker) 314

9.5 Definition (Impostor) 314

9.6 Definition (False Acceptance (Miss)) 315

9.7 Definition (False Rejection (False Alarm)) 315

9.8 Definition (Decision Tree) 331

9.9 Definition (Discrete Question) 333

9.10 Definition (Continuous Question) 333

9.11 Definition (Fixed Question) 334

9.12 Definition (Dynamic Question) 334

9.13 Definition (Binary Question) 335

10.1 Definition (Maximum Entropy (Restatement According to Good)) . . 346

10.2 Definition (Principle of minimum relative entropy (minimum discriminability)) 347

13.1 Definition (Discrete Markov Chain) 416

13.2 Definition (Unifilar Markov Source) 419

13.3 Definition (Non-Unifilar Markov Source) 420

13.4 Definition (Likelihood Ratio) 454

15.1 Definition (Vapnik-Chervonenkis Dimension – VC Dimension) 493

23.1 Definition (Identity Matrix) 635

23.2 Definition (Transpose of a Matrix) 635

23.3 Definition (Hermitian Transpose) 635

23.4 Definition (Hermitian Matrix) 636

23.5 Definition (Inverse of a Square Matrix) 636

23.6 Definition (Kronecker Product) 636

23.7 Definition (Euclidean Norm of a Vector) 636

23.8 Definition (L_p -norm of a vector) 637

23.9 Definition (Linear Dependence / Independence) 637

23.10	Definition (Euclidean (Frobenius) Norm of a Matrix)	637
23.11	Definition (Unitary / Orthogonal Matrices)	638
23.12	Definition (Conjugacy, Orthogonality, and Orthonormality)	638
23.13	Definition (Singular Values of a Matrix)	639
23.14	Definition (Rank of a Matrix)	639
23.15	Definition (Singular Value Decomposition)	639
23.16	Definition (Pseudo-Inverse (Moore-Penrose Generalized Inverse))	640
23.17	Definition (Positive Definiteness)	640
23.18	Definition (Stochastic Matrix)	643
23.19	Definition (Cone)	644
23.20	Definition (Polyhedral Cone)	644
24.1	Definition (Imaginary Number)	648
24.2	Definition (Modulus or Magnitude of a Complex Number)	648
24.3	Definition (A Circle in the Complex Plane)	650
24.4	Definition (Distance between two Complex Variables)	650
24.5	Definition (A Hermitian Function)	651
24.6	Definition (Limit of a Sequence of Numbers)	651
24.7	Definition (One Sided Limit of a Function – Right Hand Limit)	651
24.8	Definition (One Sided Limit of a Function – Left Hand Limit)	651
24.9	Definition (Limit of a Function of a Continuous Variable)	651
24.10	Definition (Positive Infinite Limit)	652
24.11	Definition (Negative Infinite Limit)	652
24.12	Definition (Continuity of Functions)	653
24.13	Definition (Discontinuous Functions)	653
24.14	Definition (A Point of Ordinary Discontinuity)	653
24.15	Definition (A Point of Infinite Discontinuity)	655
24.16	Definition (A Point of Oscillatory Discontinuity)	655
24.17	Definition (Continuity of a Function in an Interval)	657
24.18	Definition (Boundedness)	657
24.19	Definition (Continuity Class (Degree of Continuity))	657
24.20	Definition (Smoothness)	657
24.21	Definition (Piecewise Continuity)	658
24.22	Definition (Piecewise Smoothness)	658
24.23	Definition (Convex Function)	658
24.24	Definition (Strictly Convex Function)	658
24.25	Definition (Concave Function)	661
24.26	Definition (Strictly Concave Function)	661
24.27	Definition (Odd Functions)	662
24.28	Definition (Even Functions)	662
24.29	Definition (Periodic Function)	663
24.30	Definition (Periodic Extension of a Function)	663
24.31	Definition (Differentiation of Functions of Complex Variables)	663
24.32	Definition (Partial Differentiation Notation)	664
24.33	Definition (Laplace's Equation)	665

24.34	Definition (Analytic Function)	665
24.35	Definition (Pointwise Analyticity of Functions)	665
24.36	Definition (Cauchy-Riemann Conditions)	667
24.37	Definition (Harmonic Conjugate)	671
24.38	Definition (Absolutely Integrable)	673
24.39	Definition (Riemann Integral (Definite Integral))	673
24.40	Definition (Simply Connected Domain)	676
24.41	Definition (Length of a Contour Γ in \mathbb{C})	677
24.42	Definition (Taylor Series (Expansion of an analytic function into a Power Series))	683
24.43	Definition (Laurent Series (Expansion of analytic functions in an Annular Region))	684
24.44	Definition (Zeros of a Function)	685
24.45	Definition (Isolated Singularities and Poles of a Function)	686
24.46	Definition (Meromorphic Functions)	686
24.47	Definition (Finite-Domain Convolution)	688
24.48	Definition (Infinite-Domain Convolution (Convolution))	688
24.49	Definition (Inner Product of Functions)	690
24.50	Definition (Orthogonality of Functions)	690
24.51	Definition (Orthogonality of a Set of Functions)	690
24.52	Definition (Orthogonality of a Set of Functions about a Weighting Function)	691
24.53	Definition (Complete Space)	693
24.54	Definition (Linear Integral Equations)	695
24.55	Definition (General Integral Transform)	695
24.56	Definition (Kernel Function)	696
24.57	Definition (Hermitian Kernel)	696
24.58	Definition (Symmetric Kernel)	696
24.59	Definition (Gram Matrix (Kernel Matrix))	697
24.60	Definition (Definite Kernel)	697
24.61	Definition (Mercer Kernel)	697
24.62	Definition (Degenerate Kernel)	699
24.63	Definition (Dirichlet Conditions)	709
24.64	Definition (Complex Fourier Series Expansion)	709
24.65	Definition (Dirichlet Conditions with Period Normalization)	710
24.66	Definition (Complex Fourier Series Expansion with Period Normalization)	710
24.67	Definition (General Fourier Series)	712
24.68	Definition (Pointwise Convergence in an Interval)	713
24.69	Definition (Uniform Convergence in an Interval)	713
24.70	Definition (Laplace Transform)	718
24.71	Definition (Unilateral (One-Sided) Laplace Transform)	718
25.1	Definition (Global Minimzer)	774
25.2	Definition (Strict Local Minimzer (Strong Local Minimzer))	774

List of Theorems, Lemmas, and Laws

3.1	Theorem (Sampling Theorem)	79
3.2	Theorem (Extended Sampling Theorem – Fogel)	84
3.3	Theorem (Extended Sampling Theorem – Proposed here)	84
6.1	Law (De Morgan’s Law)	206
6.1	Theorem (Total Probability)	225
6.2	Theorem (Bayes Theorem)	225
6.3	Theorem (Radon-Nikodým)	229
6.4	Theorem (Hölder’s Inequality)	237
6.5	Theorem (Minkowski’s Inequality or Triangular Inequality)	237
6.6	Theorem (Jensen’s Inequality)	241
6.7	Theorem (Khintchine’s Theorem (Weak Law of Large Numbers – WLLN))	255
6.8	Theorem (Strong Law of Large Numbers (SLLN))	256
7.1	Theorem (Entropy)	271
7.2	Theorem (Asymptotic Equipartition Property (AEP))	274
23.1	Theorem (Conjugate Directions)	642
23.2	Theorem (Inverse of a Matrix)	643
23.3	Theorem (General Separation Theorem)	644
23.1	Lemma (Farkas’ Lemma)	645
24.1	Theorem (Modulus of the product of two Complex Numbers)	650
24.2	Theorem (de Moivre’s Theorem)	650
24.3	Theorem (Convex Function)	658
24.4	Theorem (Relation between existence of derivative and continuity)	665
24.5	Theorem (Cauchy-Riemann Theorem)	667
24.6	Theorem (Alternate Cauchy-Riemann Theorem)	668
24.7	Theorem (Necessary and Sufficient Cauchy-Riemann Theorem (General Analyticity))	669

24.8 Theorem (Analyticity of the Exponential Function)	671
24.9 Theorem (Analyticity of the Trigonometric Functions)	672
24.1 Lemma (Riemann's Lemma)	673
24.10 Theorem (Mean Value Theorem (Law of Mean))	676
24.11 Theorem (Cauchy Integral Theorem)	676
24.12 Theorem (Absolute Integral Bound)	678
24.13 Theorem (Cauchy Integral Formula)	679
24.14 Theorem (Morera's Theorem (Converse of Cauchy's Integral Theorem))	683
24.15 Theorem (The Cauchy Residue Theorem)	686
24.16 Theorem (The Residue Evaluation Theorem)	687
24.17 Theorem (Bessel's Inequality)	691
24.18 Theorem (Least Squares Estimation (Approximation in the Mean)) . .	692
24.19 Theorem (Completeness Relation (Bessel's Identity in a Complete Space))	693
24.20 Theorem (Schwarz Inequality for Positive Semidefinite Kernels)	697
24.21 Theorem (Hilbert's Expansion Theorem)	698
24.22 Theorem (Schmidt's Extension to Hilbert's Expansion Theorem)	698
24.23 Theorem (Mercer's Theorem)	698
24.24 Theorem (Parseval's Theorem – Fourier Series)	714
24.25 Theorem (Existence and Boundedness of the Unilateral Laplace Transform)	718
24.26 Theorem (Convolution Theorem)	726
24.27 Theorem (Correlation Theorem)	726
24.28 Theorem (Parseval's Theorem – Fourier Transform)	726
24.29 Theorem (Wiener-Khinchine Theorem)	729
24.30 Theorem (Initial Value Theorem)	758
24.31 Theorem (Final Value Theorem)	758
24.32 Theorem (Real Convolution)	759
25.1 Theorem (Powell's First Theorem)	806
25.2 Theorem (Powell's Second Theorem)	806
25.1 Lemma (No Feasible Descent Directions at the Minimizer)	826
25.2 Lemma (Feasibility Directions Subset of Linearized Feasibility Directions)	828
25.3 Lemma (Constraint Qualification – Sufficient Conditions)	828
25.3 Theorem (Kühn-Tucker Necessary Conditions – First Order Necessary Conditions for a local minimizer)	829
25.4 Theorem (Second Order Necessary Condition for a local minimizer) .	830
25.5 Theorem (Sufficient Condition for a local minimizer)	830
25.6 Theorem (Wolfe Duality Theorem)	831

List of Properties

6.1	Property (Probability of the Impossible Event)	222
6.2	Property (Probability of the Complement)	223
6.3	Property (Probability of a Union)	223
6.4	Property (Scaling)	248
6.5	Property (Translation)	248
6.6	Property (Scaling)	249
6.7	Property (Translation)	249
7.1	Property (Zero Entropy)	272
7.2	Property (Maximum Entropy)	273
7.3	Property (Averaging increases uncertainty)	276
7.4	Property (Information reduces uncertainty)	278
24.1	Property (Properties of Complex Variables)	649
24.2	Property (Triangular Inequality in the Complex Plane)	649
24.3	Property (Product of Complex Variables)	649
24.4	Property (Quotient of Complex Variables)	649
24.5	Property (Euler identities)	650
24.6	Property (Boundedness of a Continuous Function)	657
24.7	Property (Odd and Even Functions)	663
24.8	Property (Differentiation of Functions of Complex Variables)	663
24.9	Property (Properties of the Riemann Integral (Definite Integral))	673
24.10	Property (Uniqueness of Power Series)	685
24.11	Property (Addition and Multiplication of Power Series)	685
24.12	Property (Division of Power Series)	685
24.13	Property (Commutativity of Convolution)	689
24.14	Property (Associativity of Convolution)	689
24.15	Property (Distributivity of Convolution)	689
24.16	Property (Scaling Associativity of Convolution)	689
24.17	Property (Reproducing Kernel)	707
24.18	Property (Fourier Coefficients of a Real function are Real)	712

Index

- L_1 Soft Margin SVM, 501
- L_2 Soft Margin SVM, 502
- L_p
 - Distance, 303
 - Norm, 637
- L_p Distance, 237
- N -class
 - Problem, 518
- Γ -class
 - Problem, 518
- α Order Entropy, 279
- χ^2 Directed Divergence, 310
- χ^α Divergence, 310
- δ Divergence, 309
- μ -Law, 851
 - PCM, 843
- \mathcal{L}_p Class of Functions, 236
- \mathcal{L}_p Space, 236
- σ -Algebra, 214
- σ -Field, 214
- 2-Wire Telephone Channel, 532
- 2-class
 - Problems, 493
- 3G, 845
- 4-Wire Telephone Channel, 532
- 4G, 845

- A-Law, 851
- A-Law PCM, 843
- a.s.
 - Convergence, 251
- AANN, 469
- abbreviations, list of, xxvii
- Absolute
 - Integral
 - Bound, 678
- Absolute Continuity of Measures, 229

- Absolutely Integrable, 673
- acronyms, list of, xxvii
- Activation Function, 466
 - Nonlinear, 465
- Adaptation
 - MAP, 603
 - MAPLR, 607
 - MCE, 605
 - Over Time, 601
 - Speaker, 530
 - WMLLR, 607
- Adaptive
 - Differential
 - PCM, 845
 - Learning Control, 465
- Adaptive Test Normalization, 582
- Adenine, *see* DNA Adenine
- Adjoint Matrix, 636
- ADPCM, 845
- AEP, 274
- Age
 - Classification, 8
- Aging, 595
- AIC, 350
- Akaike
 - Information Criterion, *see* AIC
- Alarm, 313
- Algebra, 212
 - Boolean, 212
- Algorithm
 - Baum-Welch, 433
 - Clustering
 - Efficient k-Means, 368
 - EM, 381
 - Expectation Maximization, *see* EM
 - Fuzzy k-Means, 377
 - GEM, 387

- Generalized Expectation Maximization, *see* GEM
- Global k-Means, 368
- Hybrid, 380
- K-Harmonic Means, 378
- k-Means, 360
- k-Means Overpartitioning, 364
- k-Means Wrappers, 368
- k-Means++, 371
- Kernel k-Means, 369
- LBG, 372
- Linde-Buzo-Gray, 372
- Modified k-Means, 365, 366
- Rough k-Means, 375
- x-Means, 372
- Fast
 - Fourier Transform, 736
- FFT, 736
- Forward Pass, 430
- Forward-Backward, The, 433
- Goertzel, 749
- Goertzel's, 848
- Levinson-Durbin, 183
- Match, 430
- Schür Recursion, 183
- Viterbi, 432
- Ali-Silvey
 - Divergence, 308
- Aliasing, 99
- All Pole, 176
- All-Pole, 178, 183
- All-Zero, 178
- Allophone, 108
- Almost Everywhere, 229
- Almost Surely
 - Convergence, 251
- Alternative Hypothesis, 313
- Alveolar
 - Sounds, 117
 - Plato-Alveolar, 117
 - Post Alveolar, 118
- Alveolar Ridge, 48
- Amerindian Languages, 112
- Amplitude
 - Compression, 851
 - Finite, 655
 - Infinite, 655
 - Quantization
 - Error, 85
- Analysis
 - Factor, 404, 530
 - Latent, 530
 - FFT, 736
 - Goertzel, 749
 - LPC, 176
 - Cepstral Features, 176
 - Perceptual Linear Predictive, 190, 191
 - PLP, 190
 - Principal Component, 394
 - Nonlinear, 399
 - Spectral, 157, 168
- Analytic
 - Function, 665
- Analyticity, 665
- Exponential
 - Function, 671
- Pointwise, 665
- Trigonometric
 - Functions, 672
- Anatomy, 43
 - Auditory Cortex, 66
 - Primary, 66
 - Secondary, 66
 - Tertiary, 66
 - Auditory System, 49
 - Brain, 66
 - Ear, 49
 - Ear, 49
 - Language Production, 64
 - Language Understanding, 69
 - Neural, 65
 - Speech, 43
 - Speech Perception, 69
 - Vocal System, 44
- Angola, 123
- Annealing, 811
 - of Steel, 367
 - Process, 367
 - Simulated, 367, 474, 811
- Anti-Aliasing, 99, 153
- Aphasia, 59
 - Aconduction, 69
 - Leborgne, 60
- API, 854, 855
- Applications, 16
 - Access Control, 19
 - Audio Indexing, 19
 - Conferencing, 21
 - Financial, 16
 - Forensic, 18
 - Health, 16
 - Indexing, 19
 - Lawful Intercept, 20
 - Legal, 18
 - Other, 23
 - Proctorless Oral Testing, 21
 - Security, 19
 - Speaker Diarization, 19

- Speaker Indexing, 19
- Surveillance, 20
- Teleconferencing, 21
- Video Indexing, 19
- AR, 176, 179
- Arimoto
 - Divergence, 308
- Arithmetic
 - Mean, 257
- ARMA, 178
- Articulation, *see* Phone
- Associativity
 - Convolution, 689
 - Scaling, 689
- Assumption
 - Regularity, *see* Regularity Assumption
- Asymptotic
 - Equipartition
 - Property, 274
- Atoms of a Sample Space, 270, 282
- ATP, 52
- Audio
 - Encapsulation, 849
 - Header, 852
 - WAV, 849
 - Format, 842
 - Other, 848
 - Level, 564
 - Sampling, 77
 - Segmentation, 532
 - Volume, 564
- Audio Format
 - HE-AAC, 844
- Audio Indexing, 19
- Auditory Cortex, 49, 66, 69
 - Primary, 66
 - Secondary, 66, 68
 - Tertiary, 66, 68
- Auditory Nerve Bundle, 49
- Auditory Perception, 144
- Auditory System, 43
- Auricle, 25
- Authentication, 3
- Auto Associative
 - Neural Networks, *see* AANN
- Autocorrelation Matrix, 182
- Autoregression, 176
- Autoregressive Model, *see* AR
- Autoregressive Moving Average, *see* ARMA
- Axon, 51
 - Myelin Layer, 53
 - Myelin Sheath, 53
 - Ranvier Node, 53
 - Schwann Cell, 53
 - Sheath, 53
- Background
 - Model, 527
 - Universal, 527
- Background Model, 6
- Background Models, 528
- Bacteria
 - Pathogenic, 25
- Banach Space, 218
- Bandlimited, 79
- Bandwidth, 79
- Bark Frequency Warping, 191
- Bark Scale, 148
- Bark Warping, 169
- Bartlett Window, 163, 165
- Base
 - Countable, 207
- Basic
 - Clustering, 359
- Basis
 - Orthonormal, 396
- Baum-Welch, 433
- Bayes
 - Thomas, 225
- Bayes Theorem, 225
- Bayesian
 - Classifier, 322
 - Decision Theory, 316
 - Binary, 320
 - Information Criterion, *see* BIC
- Beckman, 798
- Bell Laboratories, 486
- Beltrami, E., 639
- Bernoulli
 - Random
 - Process, 414
 - Random Variable, 247
- Bessel
 - Function, 85
 - Identity, *see* Identity
 - Inequality, *see* Inequality
- Bessel's Inequality, *see* Inequality
- BFGS, 474, 783–785, 789
- Bhattacharyya
 - Divergence, 307
- Bhattacharyya Distance, 306
- Bibliography, 861
- BIC, 353
- Bilabial
 - Sounds, 118
- Binary
 - Channel
 - Symmetric, 284

- Hypothesis
 - Bayesian Decision, 320
 - Symmetric Channel, 284
- Biometric
 - Encryption, 625
 - Menagerie, 537
 - Privacy, 624
- Biometrics, 3, 23
 - Ear, 25
 - Face, 27
 - Finger Geometry, 30
 - Fingerprint, 28
 - Gait, 33
 - Hand Geometry, 30
 - Handwriting, 34
 - Iris, 30
 - Keystroke Recognition, 35
 - Multimodal, 26, 35
 - Palm, 28
 - Retina, 31
 - Signature, 34
 - Thermographic Imaging, 32
 - Vein, 32
- Bit
 - Rate
 - Variable, 852
- Blackman Window, 165
- Boolean Algebra, 212
- Borel σ -Field, 214
- Borel Field, 214
- Borel Sets, 214
- Bound
 - Absolute
 - Integral, 678
- Boundedness, 657
 - Continuous
 - Function, 657
- Bounds
 - Risk, 493
- Brain, 51
 - Arcuate Fasciculus, 69
 - Auditory Cortex, 66
 - Broca's Area, 59, 64
 - Brodman Areas, 57
 - Cerebrum, 54
 - Corpus Callosum, 69
 - Diencephalon, 54
 - Epithalamus, 54
 - Fissure
 - Sylvian, 66
 - Forebrain, 54
 - Frontal Lobe, 55
 - Function Localization, 58
 - Geschwind's Territory, 69
 - Hindbrain, 54
 - Hypothalamus, 54
 - Lateral Sulcus, 56, 66
 - Lobe
 - Frontal, 66
 - Occipital, 66
 - Parietal, 66
 - Temporal, 66
 - Medulla Oblongata, 54
 - Mesencephalon, 54
 - Midbrain, 54
 - Occipital Lobe, 55
 - Parietal Lobe, 55
 - Pons, 54
 - Prosencephalon, 54
 - Rhombencephalon, 54
 - Stem, 54
 - Sylvian Fissure, 56
 - Telencephalon, 54
 - Temporal Lobe, 55
 - Thalamus, 54
 - Ventriculus Tertius, 54
 - Wernicke's Area, 69
- Branch and Bound, 816
- Branches, 5
- Broca, 59
- Broca's Area, *see* Brain
- Brodman Areas, *see* Brain
- Brodman Cytoarchitectonic Areas, 57
- Broyden, 474, 783, 787
 - Family, 791
- Broyden-Fletcher-Goldfarb-Shanno, *see* BFGS
- Cadence, 33
- Canary Islands, 63, 119, 137
- Cancellation
 - Echo, 564
- Capacity
 - Learning, 350, 486, 489, 493
- Carbon Button Microphone, 613
- Carbon Microphone, 613
- Caret, 55
- Cartesian
 - Product, 216
 - Space, 232
 - Product Space, 488
- Categories, 5
- Cauchy
 - Integral
 - Formula, 679
 - Theorem, 676, 683
 - Sequence, 251
- Cauchy-Riemann
 - Conditions, 667

- Theorem, 667–669
- CDMA
 - CDMA2000, 847
 - CDMAOne, 847
- CDMA2000, 847
- CDMAOne, 847
- CELP, 847
 - CS-ACELP, 847
 - QCELP, 847
 - Qualcomm, 847
- Central Asia, 129
- Cepstra, 173
- Cepstral
 - Mean
 - Normalization, 569
 - Mean and Variance Subtraction, 569
 - Mean Subtraction, 567, 569
- Cepstral Coefficients, 173
- Cepstral Features, 173
- Cepstrum, 173, 580
- Cerebral Cortex, 55
- Cerebrum, 54
- Chain
 - Markov, 415, 643
 - Rule, 471, 664
- Chameleons, 537
- Channel
 - Binary
 - Symmetric, 284
 - Discrete, 282
 - Memoryless, 283
 - Memoryless, 432
 - Symmetric
 - Binary, 284
- Characteristic Function, 696
- Choice, 269
- Cholesky
 - Decomposition, *see* Cholesky Factorization
 - Factorization, 623, 813
- Circle
 - Complex
 - Plane, 650
- Class, 205, 207
 - \mathcal{L}_p , 236
 - Closure, 208
 - Equivalence, 208
 - Regression, 606
- Class Normalization, 582
- Classification, 3
 - Age, 8
 - Event, 8, 550
 - Gender, 8
 - Sequences, 328
 - Sex, 8
 - Speaker, 8, 550
- Classifier
 - Bayesian, 322
- Clicks
 - Consonants, 120
- Closed, 208
- Closet-Set Identification, 7
- Cloud Detection, 486
- Cluster Normalization, 582
- Clustering
 - Agglomerative, 389
 - Basic Techniques, 359
 - Bottom-Up, 389
 - Divisive, 389
 - Hierarchical, 388
 - Merging, 364
 - Supervised, 357
 - Top-Down, 389
 - Unsupervised, 341, 357, 359
- CMA, 580
- CNG, 563
- Cochlea, 145, 146
- Cochlear Fenestra Ovalis, 145
- Code
 - Convolutional, 432
- Code Excited
 - Linear Prediction, *see* CELP
- Codebook, 553
- CoDec, 842
- Coder/Decoder, *see* CoDec
- Coding
 - Difference, 845
 - Linear Predictive, 176
- Coefficient
 - Correlation, 245
 - Integral
 - Equation, 695
- Coefficients
 - Cepstral, 176
 - LAR, 176
 - Linear Predictive
 - Cepstral, 176
 - Coding, 176
 - Log Are Ratio, 176
 - LPC, 176
 - PARCOR, 176
 - Partial Correlation, 176
 - Reflection, 186
- Cohort, 6, 529
- Coin Toss
 - Example, 414
- Comfort Noise Generation, 563
- Commutativity
 - Convolution, 689

- Comparison
 - Model Performance, 453
- Compensation, 561
- Complementarity
 - Condition, 825
 - Strict, 825
- Complete
 - Metric Space, 218
 - Space, 216, 220, 693
- Completeness, 216, 218, 220
 - Relation, 694
 - Strong, 218
 - Weak, 218
- Complex
 - Function
 - Differentiation, 663
 - Number
 - Magnitude, 648
 - Modulus, 648
 - Numbers
 - Modulus of Product, 650
 - Plane
 - Circle, 650
 - Triangular Inequality, 649
 - Variables, 648
 - Distance, 650
 - Product, 649
 - Properties, 649
 - Quotient, 649
- Compression
 - Amplitude, 851
- Computer
 - Von Neumann, 465
- Concave, 773
 - Function, 661, 773
 - Strictly, 661
- Concavity
 - of Functions, 658
- Condition
 - Complementarity, *see* Complementarity
 - Condition
 - Mercer, 506
 - Quasi-Newton, 779
- Condition Number, 791
- Conditional
 - Expectation, 240
 - Expected
 - Value, 240
 - Probability, 224
- Conditional Entropy, 277
- Conditions
 - Cauchy-Riemann, 667
 - Dirichlet, 709, 710
 - Regularity, *see* Regularity Conditions
 - Wolfe-Powell, 783
- Conferencing, 21
- Confluent State, 458
- Conjugacy, 638
- Conjugate
 - Direction
 - Gradient-Free, 804
 - Directions, 800
 - Gradient, 474
 - Continuous Partan, 802
 - Fletcher-Reeves, 798
 - Iterative Partan, 802
 - Harmonic, 671
- Conjugate Duality, 831
- Conjugate Gradient, 793
- Conjugate Vectors, 638
- Consistent
 - Estimate, 253
 - Statistic, 253
- Consonant, 119
- Consonants
 - Clicks, 120
 - Nasal, 139
 - Non-Pulmonic, 120
 - Clicks, 120
 - Ejectives, 121
 - Voiced Implosives, 120
 - Pulmonic, 115
- Constant Modulus Algorithm, *see* CMA
- Constrained
 - Minimization, 820
 - Optimization, 814, 820, 835
- Constraint
 - Qualification, 827, 835
- Constraints
 - Active, 495, 822, 825
 - Equality, 815, 819–822, 833
 - Linearized, 818
 - Holonomic, 816
 - Inactive, 825
 - Inequality, 815, 816, 822, 824, 834
 - Linearized, 818
 - Linearized, 819
 - Non-Holonomic, 816
- Continuity, 652
 - Class, 657
 - Continuity, 657
 - Degree of, 657
 - Equivalent, 229
 - Function, 653
 - In an Interval, 657
 - Piecewise, 658
 - Relative Absolute
 - of Measures, 229

- Continuous
 - Entropy, 284
 - Function, 657, 665
 - Random Variable, 226, 227
 - Source, 269
- Continuous Variable
 - Limit
 - Function, 651
- Contour
 - Length, 677
- Convergence
 - a.s., 251
 - Almost Surely, 251
 - Global, 835
 - in Probability, 250, 251
 - in the Mean, 694
 - Pointwise, 713
 - Quadratic, 792
 - Random Variable, 250
 - Sampling Theorem
 - Criteria, 84
 - Sequence, 250
 - Strong, 251
 - Uniform, 713
 - Weak, 250
 - with Probability 1, 251
- Conversion
 - Intensity to Loudness, 193
- Convex, 773
 - Function, 241, 658, 773
 - Strictly, 658, 661
 - Metric
 - Space, 218
 - Pseudo, 831
 - Set, 218
 - Subset, 218
- Convexity
 - of Functions, 658
- Convolution
 - Associativity, 689
 - Commutativity, 689
 - Distributivity, 689
 - Probability Density, 234, 250
 - Scaling Associativity, 689
- Convolutional
 - Code, 432
- Coordination, *see* Phone
- Correlation, 689
 - Coefficient, 245
 - Probability Density, 234
- Cosine Kernel, *see* Kernel
- Cosine Transform
 - Discrete
 - Goertzel Algorithm, 749
- Countable Base, 207
- Countable Space, 207
- Countably Infinite, 207
- Covariance, 245
 - Matrix, 394, 395
- Covariance Function, *see* Kernel Function
- Covariance Matrix, 260
- Crisp
 - Logic, 209
- Critical Band, 169
- Critical Frequencies, 169
- Cross
 - Entropy, 290
 - Differential, 290
 - Minimum, 346
- Cross-Validation, 461, 492
 - k-Fold, 461, 492
 - Leave-One-Out, 462, 492
 - LOO, 462, 492
- CS-ACELP, 847
- Csiszár
 - Divergence, 308
- Cumulative
 - Distribution
 - Function, 235
 - Probability
 - Distribution, 247
- Curve
 - DET, 592
 - Detection Error Trade-Off, 592
 - Receiver Operating Characteristic, 590
 - Relative Operating Characteristic, 590
 - ROC, 590
- Cycle
 - Pitch, 195
- Cyrus the Great, 129
- Cytoarchitectonic Areas, 57
- Cytoarchitecture, 65
- Cytoplasm, 52
- Cytosine, *see* DNA Cytosine
- Data
 - Held-Out, 453
 - Pooling, 529
- Data Quality, 627
- dATP, *see* DNA dATP
- Davidon, 474, 783, 789, 792
 - No Line Search, 790
- Davidon-Fletcher-Powell, *see* DFP
- Davies-Swann-Campsey Method, 805
- DCT, 848
 - Goertzel
 - Algorithm, 749
- dCTP, *see* DNA dCTP

- de Moivre's Theorem, 650
- De Morgan's law, 206
- Deci Bell, *see* dB
- Decision
 - Trees, 331
- Decision Theory
 - Bayesian, 316
 - Binary, 320
- Declination, 134
- Decoding
 - HMM, 423
- Decomposition
 - Singular Value, 639
- Definite
 - Integral, 673
 - Properties, 673
- Degenerate Kernel, *see* Kernel
- Deleted
 - Estimation, 461, 492
 - Interpolation, 461, 492
- Deleted Estimation
 - Leave-One-Out, 462, 492
 - LOO, 462, 492
- Delta Cepstral Coefficient, 175
- Delta-Delta Cepstral Coefficient, 175
- Dendrite, 51
- Density
 - Normal, 326
- Density Function
 - Normal, 285
 - Uniform, 285
- Dental
 - Sounds, 118
- Deoxyribonucleic Acid, *see* DNA
- Dependence
 - Linear, 637
- Derivative
 - Radon-Nikodým, 230
- Design
 - Model, 421
- DET Curve, 592
- Detection
 - Silence, 561
 - Speaker, 11
 - Speech, 561
 - Voice Activity, 561
- Determinant, 395
- DFP, 474, 782–785
- DFT, 167
- dGTP, *see* DNA dGTP
- Diagram
 - Trellis, 428
 - Venn, 205
- Dialogic, 855
- Diarization, 19
- Diencephalon, 54
- Difference
 - Coding, 845
- Differential
 - Cross
 - Entropy, 290
 - Entropy
 - Cross, 290
 - Gaussian, 285
 - Normal, 285
 - Uniform, 285
- Differential Entropy, 284
- Differential Probability Measure, 488
- Differentiation, 663
 - Function
 - Complex, 663
 - Partial
 - Notation, 664
- Dimension
 - Vapnik-Chervonenkis, *see* VC Dimension
 - VC, *see* VC Dimension
- Diphthong, 123
- Direct
 - Search
 - Hooke-Jeeves, 804
 - Wood, 804
- Direct Method, 184
- Direct Model, 157
- Directed Divergence, *see* Divergence, 301
- Directions
 - Conjugate, 800
- Dirichlet
 - Conditions, 709
- Dirichlet Conditions, 710
- Discontinuity, 652
 - Infinite, 655
 - Ordinary, 653
 - Oscillatory, 655
- Discontinuous
 - Functions, 653
- Discourse
 - Musical, 64
- Discrete
 - Channel, 282
 - Cosine Transform
 - Goertzel Algorithm, 749
 - Expectation, 248
 - Expected
 - Value, 248
 - Fourier Transform, 731
 - FFT, 736
 - Inverse, 732
 - Parseval's Theorem, 734

- Periodicity, 734
- Power Spectral Density, 735, 739
- PSD, 735, 739
- Markov
 - Process, 268
 - Source, 268
- Random Variable, 247
 - Variance, 249
 - Variance Scaling, 249
- Discrete Fourier Transform, 167
- Discrete Source, 267
- Discrete Wavelet Transform, 194
- Discrete-Time
 - Fourier Transform, 738, 751
- Discriminability
 - Minimum
 - Principle of, 346
- Discriminant Analysis
 - Linear, 401
 - Integrated Mel, 404
- Distance, 217, 301
 - L_p , 237, 303
 - between Sequences, 302
 - between Subsets, 217
 - Bhattacharyya, 306
 - Complex
 - Variables, 650
 - Euclidean, 302
 - Hamming, 302
 - Hellinger, 304
 - Mahalanobis, 303
 - Weighted Euclidean, 303
- Distortion
 - Measure, 301
- Distribution
 - Gaussian, 285
 - Normal, 259
 - Multi-Variate, 259
 - Probability, 247
 - Cumulative, 247
- Distribution Function
 - Cumulative, 235
- Distributivity
 - Convolution, 689
- Divergence, 301
 - χ^2 , 310
 - δ , 309
 - Ali-Silvey, 308
 - Arimoto, 308
 - between Distributions, 304
 - Bhattacharyya, 307
 - Csiszár, 308
 - Directed, 288, 291, 301
 - χ^α , 310
 - Kullback-Leibler, 274, 286, 288, 291, 294, 298
- F, 308
 - General, 309
 - Jeffreys, 291, 305
 - Kullback-Leibler, 273, 274, 286, 288, 291, 294, 296–298, 305
 - Matsushita, 307
- Divergence Normalization, 583
- Divergence, The, 305
- DNA, 24, 25
 - Adenine, 25
 - Cytosine, 25
 - dATP, 25
 - dCTP, 25
 - dGTP, 25
 - dTTP, 25
 - Guanine, 25
 - Nucleotides, 24, 25
 - Recognition, 24, 25
 - Single Strand, 25
 - Strand, 24
 - Thymine, 25
 - Triphosphate, 25
- Domain
 - Multiply Connected, 676
 - Simply Connected, 676
- Dorn's Duality, 831
- Dot Product
 - Space, 220
- Doves, 537
- DTFT, 738, 751
- DTMF, 749, 848
- dTTP, *see* DNA dTTP
- Dual
 - Feasibility, 832, 834
 - Space, *see* Dual Representation
 - Variables, 497, 816, 833
- Dual Representation
 - Wolfe, 497
- Dual Tone Multi-Frequency, *see* DTMF
- Duality, 831, 833
 - Conjugate, 831
 - Dorn, 831
 - Fenchel, 831
 - Geometric, 831
 - Inference, 831
 - LP, 831
 - Self-Dual, 786
 - Wolfe, 497, 831, 834, 835
- Durbin, J., 183
- Dynamic
 - Programming, 432
- Dynamic Range, 155

- Dynamics
 - Mel Cepstra, 175
- Ear, 49
 - Auricle, 25
 - Canal, 25, 26, 49
 - Cochlear Fenestra Ovalis, 49
 - Drum, 49
 - External, 49
 - Inner, 50
 - Anterior Ampulla, 50
 - Cilia, 50
 - Cochlea, 50
 - Posterior Ampulla, 50
 - Scala Tympani, 50
 - Superior Ampulla, 50
 - Middle, 49
 - Incus, 49
 - Malleus, 49
 - Stapes, 49
 - Pinna, 25, 26
 - Recognition
 - Acoustic Method, 25
 - Visual Method, 25
- Echo
 - Cancellation, 564
- Effects
 - Nonlinear, 397
- Efficiency
 - Criterion
 - Statistical, 252
 - of Statistic, 252
- Efficient
 - Estimate, 252
 - k-Means, 368
 - Statistic, 252
- Eigenfaces, 397
- Eigenfunction, 696, 700
 - Expansion, 698
- Eigensystem
 - Decomposition, 394
- Eigenvalue, 394, 696, 700, 773
 - Problem, 394
 - Generalized, 397
- Eigenvalues
 - Degenerate, 396
 - Multiple, 396
 - Repeated, 396
- Eigenvector, 394
- Ejectives, 121
- Electret Microphone, 613
- Electrostatic Magnet, 613
- ELRA, 619
- EM, 381, 387, 479
 - MAP, 387
- EMD, 197, 198
- EMMA, 858
- Empirical Mode Decomposition, 197, *see* EMD
- Empirical Risk Minimization, 492
- Encapsulation
 - Audio, 849
 - Audio Format
 - Standard, *see* SAFE
- Encryption, 625
- Energy
 - Normalization, 565
- Enhancement
 - Signal, 199, 561
- Enrollment, 543
 - Quality Control of Utterances, 534
- Entropy, 269, 271, 272, 280, 345
 - Conditional, 277
 - Continuous, 284
 - Continuous Sources, 284
 - Cross, 290
 - Minimum, 346
 - Differential, 284
 - Gaussian, 285
 - Normal, 285
 - Uniform, 285
 - Discrete Sources, 269, 270
 - Generalized, 278
 - Generalized, 278, 279
 - Joint, 275
 - Maximum, 273
 - of order α , 279
 - Rényi, 278, 279
 - Relative, 273, 286, 288, 291
 - Minimum, 346
 - Zero, 272
- Envelope
 - Spectral, 8
- Enzyme, 52
- Epiglottal
 - Sounds, 116
- Epiglottis, 47
- Epithalamus, 54
- Equal Loudness
 - Pre-Emphasis, 192, 193
- Equal Loudness Curves, 151, 192
- Equal Loudness Pre-Emphasis, 192, 193
- Equal-Error Rate, 589
- Equality
 - Constraints, *see* Constraints
- Equation
 - Integral
 - Linear, 694, 695

- Laplace, 665
- Equivalence, 208
 - Class, 208
 - Relation, 208
- Equivalent, 229
- Ergodic
 - Source, 269
- Ergodicity, 269
- Error
 - Amplitude
 - Quantization, 85
 - Truncation, 102
- Error Correcting Code, 518
- Estimate
 - Consistent, 253
 - Efficient, 252
 - Periodogram, 735
- Estimation
 - Audio
 - Volume, 564
 - Deleted, 461, 492
 - Held-Out, 453
 - Maximum
 - Mutual Information, 348
 - of the Mean, 253
 - of the Variance, 258
 - Parameter, 341
 - Power
 - Spectral Density, 735
 - Power Spectral Density, 739
 - PSD, 735, 739
 - Periodogram, 735, 739
- Estimator
 - Biased, 258
- Euclidean
 - Norm, 470
 - Matrix, 637
 - Vector, 636
- Euclidean Distance, 302
- Euler
 - Identities, 650
- Euler-Lagrange, 829
- Euler-Lagrange Equations, 816
- European Language Resources Association, 619
- Evaluation
 - Results, 589
- Even
 - Function, 662
 - Properties, 663
- Event, 205
 - Blast, 8
 - Classification, 8
 - Gun Shot, 8
 - Horn, 8
 - Music, 8
 - Scream, 8
 - Whistle, 8
- Event Classification, 8, 550
- Evolution of Languages, *see* Languages
- Excess Kurtosis, 246
- Exclusive OR, 481
- Expansion
 - Series
 - Laurent, 683
 - Power, 683
 - Taylor, 683
- Expansion Theorem
 - Hilbert, *see* Theorem
 - Mercer, *see* Theorem
 - Schmidt, *see* Theorem
- Expectation, 239
 - Conditional, 240
 - Discrete, 248
 - Function, 248
 - Maximization, *see* EM
 - Generalized, *see* GEM
- Expected
 - Log
 - Likelihood, 454
 - Value, 239
 - Conditional, 240
 - Discrete, 248
 - Function, 248
- Extended
 - Function
 - Real Valued, 228
 - Real Valued
 - Function, 228
 - Sampling Theorem, 84
- Extended Real Number, 215
- Extensible
 - Multimodal Annotation, *see* EMMA
- Extensions
 - Sampling Theorem, 84
- Extraction
 - Feature, 143
- F-Divergence, 308
- FA, 404
- Factor
 - Analysis
 - Joint, 531
 - Factor, 733
 - Total, 532
- Factor Analysis, 404, 530
- Factorization, 791
 - Cholesky, 623, 813

- False Acceptance Rate, *see* FAR
- False Alarm Probability, 315
- False Rejection Rate, *see* FRR
- FAR, 315
- Farkas
 - Lemma, 825, 829
- Farkas'
 - Lemma, 645
- Farsi, *see* Languages, Persian
- Fast
 - Fourier Transform, 736
- Fast Fourier Transform, 167
- Fault Detection, 486
- Feasibility
 - Direction, 826
 - Points, 815
 - Region, 815, 816, 822, 825
- Feasible
 - Directions, 818–820, 823
 - Point, 822, 828
- Feature
 - Extraction, 143, 157
 - Jitter, 8
 - Mapping, 578
 - Shimmer, 8
 - Vector, 394, 469
 - Warping, 576
- Features
 - Cepstral, 133
 - Mel Cepstral, 173
 - Metrical, 138
 - Modulation, 197
 - Other, 193
 - Prosodic, 132
 - Suprasegmental, 107
 - Temporal, 131
 - Vocal Source, 195
 - Wavelet Filterbanks, 194
- Feedforward
 - Learning, 470, 473
 - Neural Networks, 466
 - Training, 470, 473
- Feedforward Neural Network, *see* FFNN
- Fenchel's Duality, 831
- FEPSTRUM, 197
- FFNN, 477
- FFT, 167, 736
- Field, 212, 213
 - σ , 214
 - Borel, 214
 - Lebesgue, 215
- Filter
 - Anti-Aliasing, 153
 - Hi-Pass, 153
 - Hi-pass, 176
 - High-Pass, 153
 - Kalman, 579
 - Notch, 579, 580
 - Wiener, 580
- Filtering
 - J-RASTA, 571
 - RASTA, 571
 - Relative Spectral (J-RASTA), 571
 - Relative Spectral (RASTA), 571
 - Square-root, *see* Cholesky Factorization
- Finance, 465
- Finite
 - Amplitude, 655
- Fisher
 - Information, 294
 - Matrix, 343
 - Score, 342
- Flanagan, J.L., 129
- Flatness
 - Spectral, 168
- Fletcher, 474, 783, 798, 808
- Fletcher, R., 816
- Fletcher-Reeves, 798
- Folding
 - Frequency, 87
- Forebrain, 54
- Formant, 8
- Format
 - Audio, 842
 - Other, 848
- Formats
 - Audio
 - Encapsulation, 849
- Fortmann, T. E., 399
- Forward Pass Algorithm, 430
- Forward-Backward Algorithm, The, 433
- Four-Wire Telephone Channel, 532
- Fourier, 722
 - Coefficients
 - Real, 712
 - Complex Expansion, 82
 - Generic Descriptor, *see* GFD
 - Series
 - Complex, 95
 - Transform, 580
 - Complex, 80, 98
- Fourier Transform
 - Discrete, 731
 - Parseval's Theorem, 734
 - Periodicity, 734
 - Discrete-Time, 738, 751
 - Fast, 736
 - Inverse Discrete, 732

- Fourth
 - Moment, 246
- FOXP2 Gene, 65
- Frame, 160
 - Overlapping, 176
- Framing, 160
- Frequencies
 - Fundamental, 195
- Frequency
 - Folding, 87
 - Fundamental, 8
 - Nyquist Critical, 87
- Frequency Warping, 169
- Fricative, 8
- Frobenius Norm, 637
- FRR, 315
- Full Rank
 - see Rank, 638
- Function
 - Analytic, 665
 - Bandlimited, 79
 - Bessel, 85
 - Complex
 - Differentiation, 663
 - Concave, 773
 - Continuity, 653
 - In an Interval, 657
 - Convex, 773
 - Cumulative
 - Distribution, 235
 - Discontinuous, 653
 - Even, 662
 - Exponential
 - Analyticity, 671
 - Hermitian, 651
 - Holomorphic, 665
 - Kernel, *see* Kernel Function
 - Lagrangian, *see* Lagrangian
 - Limit
 - One-Sided: Continuous, 651
 - One-Sided: Left-Hand, 651
 - One-Sided: Right-Hand, 651
 - Logistic, *see* Logistic Function
 - Loss, 318
 - Normal
 - Density, 285
 - Objective, 474
 - Odd, 662
 - One-Sided Limit
 - Continuous, 651
 - Left-Hand, 651
 - Right-Hand, 651
 - Orthogonality, 690
 - Penalty, 318
 - Periodic, 95, 663
 - Poles of, 686
 - Probability
 - Mass, 247
 - Probability Density, 229, 231
 - Joint, 233
 - Marginal, 233
 - Quadratic, 773
 - Random Variable
 - Discrete, 248
 - Regular, 665
 - Smooth, 814
 - Trigonometric
 - Analyticity, 672
 - Uniform
 - Density, 285
 - Zeros of, 685
- Function Spaces, 236
- Functions
 - Analyticity
 - Pointwise, 665
 - Concave, 658, 661
 - Strictly, 661
 - Continuous, 657, 665
 - Convex, 241, 658
 - Strictly, 658, 661
 - Inner Product, 690
 - Orthogonal Set, 690, 712
 - Orthonormal Set, 691
 - Pointwise
 - Analyticity, 665
 - Probability, 228
 - Smooth, 773
- Fundamental
 - Frequencies, 195
- Fundamental Frequency, 8
- Fundamental Sequence, 251
- Fusion, 15, 26, 35
- fusion, 612
- Fuzzy
 - k-Means, 377
 - Set
 - Theory, 211
- G.721, 845
- G.723, 845
- G.726, 845
- G.729, 847
- Gabor Transform, 167
- Gall, F.J., 58, 59
- Gauss Window, 167
- Gaussian
 - Distribution, 285
 - Multi-Dimensional, 326

- k-Means, 365
- Mixture Models, 442
- Radial Basis Function, *see* GRBF
- Gaussian Radial Basis Function
 - Kernel, *see* Kernel
- Gaussianization
 - Short-Time, 578
- GEM, 387
- Gender
 - Classification, 8
- Gender Normalization, 582
- Gene, 52
 - FOXP2, 65
- General
 - Divergence, 309
- Generalized
 - Eigenvalue Problem, 397
 - Entropy, 279
- Generalized Inverse
 - Moore-Penrose, *see* Pseudo-Inverse
- Generalized Linear Discriminant Sequence
 - Kernel, *see* GLDS Kernel
- Generic Fourier Descriptor, *see* GFD
- Geometric
 - Mean, 257
- Geometric Duality, 831
- Geschwind's Territory, *see* Brain
- GFD, 26
- GLDS Kernel, *see* Kernel
- Global
 - Convergence, 835
 - k-Means, 368
 - Minimizer, 774
 - Minimum, 475
 - Optimum, 773
 - Solution, 474
 - System
 - for Mobile Communications, *see* GSM
- Glottal
 - Sounds, 116
- Glottis, 47
- GMM, 465
 - Mixture
 - Coefficients, 460
 - Practical Issues, 451
 - Training, 451
- Goats, 536
- Goertzel
 - Algorithm, 749
- Goertzel's
 - Algorithm, 848
- Goldfarb, 783
- Golgi, C., 52, 53
- Gradient
 - Conjugate, 474, 793
 - Vector, 473
- Gradient-Free
 - Conjugate Direction, 804
 - Davies-Swann-Campey, 805
 - Powell, 805
 - Rosenbrock, 804
 - Optimization, 803
- Gram Matrix, *see* Matrix
- Gram-Schmidt Orthogonalization, 641
 - Modified, 641
- Gray Matter, 54
- Gray's Anatomy, 44, 57
- GRBF, 470
- Greenstadt, 474
- Groupe Spécial Mobile, *see* GSM
- GSM, 185, 845
 - 06.10, 845
 - 06.60, 845
 - EFR, 845
 - Enhanced
 - Full Rate, *see* GSM EFR
 - Full rate, 845
 - Enhanced, 845
 - Half Rate, 845
- GSMA, 847
- Guanine, *see* DNA Guanine
- Guyon, Isabelle, 486
- Gyri, 54
- Gyrus, 54
- Hölder's Inequality, *see* Inequality
- Half Total Error Rate, 590
- Hamilton's Principle, 816
- Hamiltonian, 816
- Hamming
 - Distance, 302
 - Window, 580
- Hamming Window, 162
- Handset Normalization, 582
- Handset Test Normalization, 582
- Handwriting
 - Recognition, 486
- Handwriting Recognition, 397
- Hann Window, 163
- Hard Palate, 48
- Harmonic, 8
 - Conjugate, 671
 - Higher, 8
 - Mean, 257
- Hawaiian, *see* Languages
- HE-AAC, 844
- Header

- Audio
 - Encapsulation, 852
- Hearing, 49, 66
 - Psychophysical Power Law, 193
- Held-Out
 - Data, 453
 - Estimation, 453, 456, 461
- Hellinger Distance, 304
- Hermansky, H., 193
- Hermitian
 - Function, 651
 - Kernel, *see* Kernel
- Hermitian Matrix, 636
- Hermitian Transpose, 635
- Hessian, 773, 787
 - Inverse, 787
 - Matrix, 473, 829
- Hestenes, 798
- Hi-Pass Filter, 153
- Hi-pass Filter, 176
- Hidden
 - Markov
 - Model, 411, 418, 643
 - Toolkit, *see* HTK
- Hierarchical
 - Mixture of Experts, *see* HME
 - Mixtures of Experts, *see* HME
- High-Pass Filter, 153
- Higher Order
 - Statistics, 381
- Hilbert Space, 220
- Hilbert's Expansion Theorem, *see* Theorem
- Hilbert, David, 486
- Hindbrain, 54
- Histogram
 - Equalization, 570
- HME, 465, 479
- HMM, 411, 418, 465, 479, 480
 - Decoding, 423
 - Practical Issues, 451
 - Training, 423, 451
- Holmgren, E., 700
- Holomorphic
 - Function, 665
- Holonomic
 - Constraints, 816
- Homogeneous
 - Objective Function, 787
- Homomorphic
 - Deconvolution, 766
 - System Theory, 763
- Hooke-Jeeves
 - Direct Search, 804
- Hoshino, 474, 789
- Hoshino's Method, 789
- HTER, 590
- HTK, 628
- Hybrid
 - Clustering
 - Algorithm, 380
- Hybridization Process, 24, 25
- Hypothalamus, 54
- Hypothesis
 - Alternative, 313
 - Null, 313
 - Testing, 313
- Hypothetical Menagerie, 536
- i-Vector, 532
- i-Vectors, 532
- i.i.d. Random Variables, 255, 256
- IBM, 855
- Ideal Sampler, *see* Sampler
- Identification, 3, 548
 - Closed-Set, 7
 - Open-Set, 7
- Identification Results, 593
- Identity
 - Bessel, 693
- Identity Matrix, 635
- Illumination Variations, 26
- Image
 - Processing, 465
 - Recognition, 486
- Image Recognition, 397
- Imaginary
 - Number, 648
- IMELDA, 404
- IMF, 198
- Impedance
 - Specific Acoustic, 149, 150
- Implosives
 - Voiced, 120
- Impostor, 314
- Incomplete
 - Data
 - Estimation, 381
- Independence
 - Linear, 637
 - Statistical, 226
- Index
 - Notation, 471, 472
- Indexing, 19
 - Audio, 19
 - Speaker, 19
 - Video, 19
- Indiscernibility
 - Relation, 208

- Individual
 - Speaker Model, 526
- Inequality
 - Bessel, 691, 692, 694
 - Constraints, *see* Constraints, *see* Constraints
 - Hölder, 236, 237, 242, 690
 - Jensen, 241, 242, 248
 - Minkowski, 236, 237
 - Schwarz, 237, 242, 244, 690, 697
 - Triangular, 217, 237
- Inexact
 - Line Search, 789
- Inference Duality, 831
- Infinite
 - Amplitude, 655
 - Countably, 207
 - Dimensional Space, 219
 - Discontinuity, 655
- Information, 269, 279, 280
 - Criterion
 - Akaike, *see* AIC
 - Bayesian, *see* BIC
 - Residual, *see* RIC
 - Discrete Sources, 279
 - Fisher, 294, 343
 - Loss, 104
 - Mutual, 291
 - Relative, 321
- Information Source, 266
- Information Theory, 265
- Initial Scaling Quasi-Newton, 788
- Initiation, *see* Phone
- Inner Product
 - of Functions, 690
 - Space, 220
- Inner Product Space, 219
- Inseparable
 - Linearly, 500
- Integer
 - Programming, 816
- Integrable
 - Absolutely, 673
- Integral, 227
 - Bound
 - Absolute, 678
 - Definite, 673
 - Equation
 - Coefficient, 695
 - Linear, 694, 695
- Formula
 - Cauchy, 679
 - Riemann, 673
- Theorem
 - Cauchy, 676
 - Transform
 - General, 695
- Integral Equation
 - Linear
 - Second Kind, 486
- Integral Transform, 647
- Integrated Mel Linear Discriminant Analysis, 404
- Integration, 672
 - Measure, 227
- Intensity, 149, 172
 - Relative, 150
- Intensity to Loudness Conversion, 193
- International Phonetic Alphabet, *see* IPA
- Internet, 509
- Interpolation
 - Deleted, 461, 492
- Intonation, 132
- Intrinsic Mode Function, 198
- Inverse, 636
 - Discrete
 - Fourier Transform, 732
 - Hessian, 787
- Image
 - Transformations, 238
- Kernel, 695
- IPA, 112
- Isolated
 - Singularities, 686
- Issues
 - Practical
 - GMM, 451
 - HMM, 451
 - Neural Networks, 479
- ITT, 855
- ITU-T, 842
- Jackel, Larry, 486
- Jacobian
 - Matrix, 791, 815, 820, 834
- Jaw, 64
- Jeffreys, 291
- Jeffreys Divergence, *see* Divergence, 305
- Jensen's Inequality, *see* Inequality
- Jerri, A.J., 85
- JFA, 450, 583
- Jitter, 8
 - Macro, 103
 - Micro, 103
- Joint
 - Factor
 - Analysis, 531
- Joint Entropy, 275
- Joint Factor Analysis, *see* JFA

- Joint Probability Density Function, 233
- Jordan, M. Camille, 639
- k-d Tree, 368, 375
- k-dimensional
 - Tree, *see* k-d tree
- K-Harmonic Means Algorithm, 378
- K-Means, 360
- k-Means, 365
 - Efficient, 368
 - Fuzzy, 377
 - Gaussian, 365
 - Global, 368
 - Kernel, *see* Kernel k-Means
 - Modified, 366, 367, 811
 - Overpartitioning, 364
 - Rough, 375
 - Wrappers, 368
- k-Means++, 371
- Kühn-Tucker
 - Conditions, 829
 - Constraint
 - Qualification, *see* Constraint Qualification
 - Point, 829, 830
 - Regularity Assumption, *see* Regularity Assumption
- Kalman Filter, 579
- Karhunen Loève Transformation, *see* KLT
- Karhunen-Loève Transformation, 394
- Karush-Kühn-Tucker, *see* -Kühn-Tucker
- Kernel, 85, 695
 - Cosine, 508
 - Definite, 697
 - Degenerate, 699, 701
 - Eigenfunction, 700
 - Eigenvalue, 700
 - Expansion, 698
 - Function, 485, 696
 - Fuzzy tanh, 512
 - Fuzzy Hyperbolic Tangent, 512
 - Gaussian Radial Basis Function, 507
 - GLDS, 509
 - Hermitian, 696
 - Indefinite, 697
 - Inverse, 695
 - Jeffreys Divergence, 511
 - k-Means, 369, 695
 - Kullback-Leibler Divergence, 511
 - Linear, 506
 - Mapping, 503
 - Mercer, 506
 - Negative Semidefinite, 697
 - Neural Network, 513
 - Non Positive Semi-Definite, 511
 - Non-Positive Semidefinite, 511, 512
 - Normalization, 513
 - PCA, 400, 514, 695
 - Positive Semidefinite, 697
 - Radial Basis Function, 507
 - Reproducing, 707
 - Symmetric, 696, 698
 - Trick, 504, 506
- Khoisan, *see* Languages
- KING, 617
- KLT, 397
- Knowledge-Based, 15, 548
- Kotel'nikov, V.A., 80
- Kramer, H.P., 85
- Kronecker
 - Product, 636
- Kullback, 291
- Kullback-Leibler
 - Kernel, *see* Kernel
- Kullback-Leibler Divergence, *see* Divergence, 305
- Kurtosis, 246
 - Excess, 246
- La Gomera, 63, 119, 137
- Labiodental
 - Sounds, 118
- Ladefoged, P., 136
- Lagrange
 - Euler-Lagrange Equations, 816
 - Multiplier, 816, 817, 820–822
 - Multipliers, 497, 825, 831, 834
- Lagrangian, 817, 821, 822, 829, 831, 833
- Lambs, 536
- Landau Asymptotic Notation, 826
- Language
 - Modeling, 411, 477
 - Production, 64
 - Recognition, 411
 - Silbo, 63, 137
 - Understanding, 69, 411
 - Whistled
 - Silbo Gomero, 119
- Languages
 - Amerindian, *see* Amerindian Languages
 - Arabic, 129
 - Canary Islands, 137
 - Chinese
 - Cantonese, 136
 - Mandarin, 135
 - Dravidian, 139
 - Dutch, 139
 - Evolution, 129
 - French, 116, 129

- German, 129
- Hausa, 120
- Hawaiian, 122
- Indian, 139
- Indo-European, 207
- Japanese, 139
- Khoisan, 123
- Margi, 120
- Persian, 116, 124, 129, 130, 135, 139, 140
 - Modern, 129
 - Old, 129
- Rotokas, 122
- Silbo Gomero, 137
- South African, 120
- Spanish
 - Silbo, 137
- Telugu, 139
- Tonal, 132
- Turkish, 138, 139
- Zulu, 120
- Laplace, 717
 - Equation, 665
 - Inversion, 720
- LAR, 176, 189
- LAR Coefficients, 176
- Large
 - Numbers
 - Law of, 254
 - Strong Law of, 255, 256
 - Weak Law of, 255
- Large-Scale
 - Optimization, 810
- Large-Scale Systems, 628
- Laryngealization, 109
- Larynx, 44, 64
- Latent
 - Factor Analysis, 530
- Latent Factor Analysis, *see* LFA
- Laurent
 - Series, 684
- Laver, J., 47, 108, 122
- Law
 - De Morgan, 206
 - Large Numbers, 254
 - Strong, 255
 - of Large Numbers, 254
 - Strong, 256
 - Weak, 255
- LBG, 372
- LDA, 401
- Learning, 341, 357
 - Capacity, 493
 - Neural Network, 473
 - Semi-Supervised, 390
- Learning Control
 - Adaptive, 465
 - Iterative, 465
- Leave-One-Out Cross-Validation, 462, 492
- Lebesgue
 - Field, 215
 - Measurable Space, 216
 - Measurable Subsets, 215
 - Measure, 215
 - Space, 236
- Leibler, 291
- Lemma
 - Farkas, *see* Farkas' Lemma
 - Riemann, 673
- Length
 - Contour, 677
- Levinson, N., 183
- Levinson-Durbin, 193
- Levinson-Durbin Algorithm, 183
- LFA, 583
- Likelihood, 288
 - Estimation
 - Maximum, *see* MLE
 - Expected, 454
 - Maximum, 381
 - Unit, 453
- Limit
 - Function
 - One-Sided: Continuous, 651
 - One-Sided: Left-Hand, 651
 - Infinite
 - Negative, 652
 - Positive, 652
 - Negative
 - Infinite, 652
 - One-Sided Function
 - Right-Hand, 651
 - Positive Infinite, 652
 - Sequence, 651
- Limits, 651
- Linde-Buzo-Gray, 372
- Line Search, 776, 778, 779, 783, 786, 787, 791, 792, 809
 - Exact, 474
 - Inexact, 474, 784, 789
 - No, 789, 790
- Line Search Free Learning, *see* LSFL
- Linear
 - Equations, 643
 - Optimization, 773
 - PCM, 842
 - Prediction
 - Code Excited, *see* CELP
 - Predictive

- Coding, *see* LPC, *see* LPC
- Perceptual, 190, 191
- Pulse Code Modulation, *see* Linear PCM
- Regression, 606
- Separability, 493
- Transformation, 403
- Linear Dependence, 637
- Linear Discriminant Analysis, 401
- Linear Independence, 637
- Linear Integral Equation
 - Second Kind, 486
- Linear Predictive Cepstral Coefficients, 176
- Linear Predictive Coding, *see* LPC, 176
- Linearly
 - Inseparable, 500
- Linguistic
 - Pitch Range, 132
 - Stress, 135
- Linguistic Temporal Features, 140
- Linguistics, 107
- Lips, 48, 64
- Liveness, 548
- LLR, 321
- Local
 - Minimizer
 - Strict, 774
 - Minimum, 475
- Log
 - Likelihood
 - Expected, 454
- Log Are Ratio Coefficients, 176
- Log Area Ratio, 176, 189
- Log Short-Term Memory, *see* LSTM
- Log-Likelihood, 288
 - Ratio, 288, 321
- Logic
 - Crisp, 209
- Logical And, 205
- Logistic Function, 472, 513
- Long
 - Short-Term Memory, 477
- LOO-Cross-Validation, *see* Leave-One-Out
 - Cross-Validation
- Loss
 - Information, 104
- Loss Function, 318
- Loudness, 137, 144, 149, 150, 172, 848
- LP Duality, 831
- LPC, 157, 176, 847, 848
- LPCC, 176
- LPCM, 842, 851
- LSFL, 475
- LSTM, 477
- MA, 157
- Maddieson, I., 122
- Magnetic Resonance Image, *see* MRI
- Magnitude
 - Complex Number, 648
- Magnitude Warping, 172
- Mahalanobis Distance, 303
- Makhoul, J., 181
- MAP, 344
 - Adaptation, 603
 - using EM, 387
 - World, 583
- MAPLR
 - Adaptation, 607
- Mapping
 - Feature, 578
 - Kernel, 503
- Marginal Probability Density Function, 233
- Markoff, *see* Markov
- Markov
 - Chain, 415, 643
 - Model, 643
 - Process, 268
- Mass Function
 - Probability, 247
- Match Algorithm, 430
- Matrices
 - Orthogonal, 638
 - Unitary, 638
- Matrix
 - Adjoint, 636
 - Autocorrelation, 182
 - Covariance, 394, 395
 - Euclidean Norm, 637
 - Frobenius Norm, 637
 - Full Rank
 - see* Rank, 638
 - Gram, 697
 - Hermitian, 636
 - Hermitian Transpose, 635
 - Identity, 635
 - Inverse, 636
 - Inversion
 - Sherman-Morrison, 642
 - Jacobian, 791
 - Norm, 636, 637
 - Positive Definiteness, 640
 - Pseudo-Inverse, *see* Pseudo-Inverse
 - Rank, 639
 - Scatter, 402
 - Singular Values, 639
 - Stochastic, 643
 - Toeplitz, 183
 - Transpose, 635

- Matsushita
 - Divergence, 307
- Maximum
 - A-Posteriori
 - Estimation, *see* MAP
 - a-Posteriori
 - Adaptation, *see* MAP Adaptation
 - Entropy, 345
 - Principle of, 346
 - Likelihood
 - Estimation, *see* MLE
 - Linear Regression, 606
 - Techniques, 381
 - Mutual Information
 - Estimation, 348
- Maximum Entropy, 273
- MCE
 - Adaptation, 605
- McLauren
 - Series, 684
- Mean, 239, 240, 257, 259
 - Arithmetic, 257
 - Discrete, 248
 - Estimation, 253
 - Geometric, 257
 - Harmonic, 257
 - Quadratic, 257
 - Sample, 259
 - Value
 - Theorem, 669, 670, 676
- Measurable
 - Space, 209
 - Subsets, 205
 - Transformations, 238
- Measurable Space, 214
 - Product, 216
- Measure, 205, 207, 209, 211, 212, 215
 - Integration, 227
 - Lebesgue, 215
 - Probability, 221
- Measure Theory, 205, 211
- Measures
 - Relative Absolute Continuity, 229
- Medulla Oblongata, 54
- Mel
 - Cepstral
 - Dynamics, 175
 - Features, 173
- Mel Cepstrum Modulation Spectrum, *see* MCMS
- Mel Frequency Cepstral Coefficients, 173
- Mel Scale, 146, 147, 169
- Mel Warping, 169
- Mel-Frequency Discrete Wavelet Coefficient, 194
- Melody Scale, *see* Mel Scale, 169
- Memory
 - Long Short-Term, 477
- Memoryless
 - Channel, 432
 - Model, 413
 - Source, 267, 413
- Menagerie
 - Biometric, 537
 - Hypothetical, 536
 - Speaker, 536
- Mercer
 - Condition, 506
 - Kernel, 506
- Mercer's Expansion Theorem, *see* Theorem
- Mercer, J., 697
- Merging
 - of Clusters, 364
- Mesencephalon, 54
- Method
 - Davies-Swann-Campey, 805
 - Hoshino, 789
 - Powell, 805
 - Rosenbrock, 804
- Methods
 - Search, 804
 - Variable Metric, 779
- Metric, 301
- Metric Space, 217
 - Complete, 218
- Metrical Features, 138
- MFCC, 173
- MFDWC, 194
- Microbolometer, 32
- Microphone, 26, 580
 - Carbon, 613
 - Electret, 613
- Microphones, 613
- Midbrain, 54
- Minimization, 473, 789
- Minimizer
 - Global, 774
 - Local Strict, 774
- Minimum
 - Cross
 - Entropy, 346
 - Discriminability
 - Principle of, 346
 - Global, 475
 - Minimum, 475
 - Relative
 - Entropy, 346

- Tolerated
 - Ratio, 396
- Minkowski's Inequality, *see* Inequality
- Miss Probability, 315
- Missing
 - Data
 - Estimation, 381
- MIT, 580
- Mixture
 - Coefficients
 - GMM, 460
 - Models
 - Gaussian, 442
- MLE, 296, 342
- MLLR, 606
- MMIE, 348
- Modalities, 12
- Model, 421
 - Background, 6, 527, 528
 - Cohort, 6
 - Hidden Markov, 411
 - Markov, 643
 - Hidden, 643
 - Memoryless, 413
 - Performance Comparison, 453
 - Quality, 534
 - Selection, 349
 - Speaker
 - Independent, 529
 - Individual, 526
 - Tractability, 449
 - Universal Background, 6
- Model Quality, 534
- Modeling
 - Hidden Markov, 411
 - Language, 411, 477
 - Speaker, 43, 525
- Modified
 - k-Means, 366, 367, 811
 - x-Means, 372
- Modulo of a Measure, 229
- Modulus
 - Complex Number, 648
 - Complex Numbers
 - Product, 650
- Moment
 - Fourth, 246
 - Second, 242
 - Statistical
 - Estimation, 253
 - First, 248
 - Third, 245
- Moments
 - First, 239, 240
 - Statistical, 239
- Moore-Penrose
 - Generalized Inverse, *see* Pseudo-Inverse
- Mora, 133
- Morera's Theorem, 683
- Morphology, 107, 129
- Motor Control, 64
- Motorola, 855
- Moving Average, *see* MA
- Moving Picture Experts Group, *see* MPEG
- MP3, 843
- MPEG, 843
- MPEG-1, 843
- MPEG-2, 843
- MPEG-4, 843
- MRI, 55
- Mu-Law
 - PCM, 843
- Multi-Dimensional
 - Gaussian
 - Distribution, 326
- Multi-Layer
 - Neural Network, 465
- Multidimensional
 - Space, 232
- Multimodal
 - Annotation
 - Extensible, *see* EMMA
 - Biometrics, 26, 35
- Multiplier
 - Lagrange, 816
- Multiply Connected
 - Domain, 676
- Musical Discourse, 64
- Mutual
 - Statistical Independence, 226
- Mutual Information, 291
 - Estimation
 - Maximum, 348
- Myelin Layer, 53
- Myelin Sheath, 53
- Namibia, 123
- NAP, 450, 583
- Narrowband Spectrogram, 88
- Nasal
 - Cavity, 47
 - System, 47
- National
 - Institute of Standards and Technology, 850
- Natural Language Understanding, *see* NLU
- Necessary Conditions
 - Local Minimizer, 829
- Nerve

- Auditory Bundle, 49
- Ending
 - Postsynaptic, 53
 - Presynaptic, 53
- Vestibulocochlear Bundle, 49
- Nervous System, 51
- Network
 - Feedforward, 466
 - Neural, 465
 - Time-Delay, 477
- Neural
 - Anatomy, 65
- Neural Network
 - Feedforward, 466, *see* FFNN
 - Learning, 473
 - Multi-Layer, 465
 - Perceptron, 466
 - Recurrent, 477
 - Time-Delay, *see* TDNN, 477
 - Training, 473
- Neural Networks, 465
 - Architecture, 465
 - Auto Associative, *see* AANN
 - Hierarchical Mixtures of Experts, *see* HME
 - HME, 479
 - Kernel, *see* Kernel
 - Practical Issues, 479
 - Radial Basis Function, *see* RBFN
 - Recurrent, 476
 - TDNN, *see* TDNN
- Neuron, 51
 - Axon, 51, 52
 - Myelin Layer, 53
 - Myelin Sheath, 53
 - Ranvier Node, 53
 - Schwann Cell, 53
 - Sheath, 53
 - Terminal Buttons, 53
 - Dendrite, 51
 - Dendrites, 52
 - Gogli Apparatuses, 52
 - Mitochondria, 52
 - Myelin Layer, 53
 - Nissl Granules, 52
 - Nucleolus, 52
 - Perikaryon, 51
 - Ribosome, 52
 - Soma, 53
 - Synapse, 51
 - Synaptic Cleft, 53
 - Terminal Buttons, 53
 - Type I, 53
 - Type II, 53
 - Vesicles, 52
- Neurotransmitter, 53
 - Serotonin, 53
- Newton
 - Minimization, 777
 - Projected, 802
 - Quasi, 474
- Newton-Raphson, 474
- NIST, 580, 850
- NN, *see* Neural Networks
- No Line Search, 789, 790
- Noise
 - Compensation, 561
 - Generation
 - Comfort, 563
 - Reduction
 - Narrowband, 579
- Nomenclature, xxxi
- Non Positive Semi-Definite
 - Kernel, 511
- Non-Holonomic
 - Constraints, 816
- Non-Pulmonic
 - Consonants, 120
 - Clicks, 120
 - Ejectives, 121
 - Implosives
 - Voiced, 120
 - Voiced
 - Implosives, 120
- Non-Stationary Signal, 76
- non-streaming, 851
- Non-Unifilar, 268, 416, 418
 - Source, 420
- Nonlinear
 - Activation Function, 465
 - Effects, 397
 - Optimization, 773
 - PCA, 399, 469
- Nonsmooth
 - Optimization, 814
- Norm, 636
 - Euclidean, 470
 - Matrix, 637
 - Vector, 636
- Normal
 - Density, 326
 - Density Function, 285
 - Vector, 815, 816
 - Vectors, 817, 820, 828
- Normalization, 576, 581
 - AT-Norm, 582
 - C-Norm, 582
 - Cepstral
 - Histogram, 570

- Mean and Variance Subtraction, 569
- Mean Subtraction, 569
- Cepstral Mean Subtraction, 567
- D-Norm, 583
- Energy, 565
- F-Norm, 583
- F-Ratio, 583
- H-Norm, 582
- HT-Norm, 582
- MAP
 - World, 583
- Speaker, 581
- T-Norm, 582
- Test Norm, 582
- Vocal Tract Length, 573
- VTLN, 573
- World MAP, 583
- World Maximum A-Posteriori, 583
- Z-Norm, 581
- Normed Vector Space, 218
- Notation
 - Index, 471, 472
- Notch Filter, 579, 580
- Novell, 855
- NTIMIT, 617
- Nucleotides, *see* DNA Nucleotides
- Nuisance Attribute Projection, *see* NAP
- Null Hypothesis, 313
- Number
 - Imaginary, 648
- Numerical
 - Stability, 813
- Numerical Stability, 623
- Nyquist
 - Critical
 - Angular Frequency, 99
 - Critical Frequency, 79, 87
 - Rate, 79
- Nyquist, H., 79
- Objective
 - Function, 474
- Objective Function
 - Homogeneous, 787
- Odd
 - Function, 662
 - Properties, 663
- OGG, 852
 - Speex, 847
 - Vorbis, 844, 847
- One-to-one
 - Transformation, 238
- Online
 - Handwriting Recognition, 397
- Open-Set Identification, 7
- Optimization, 473
 - Constrained, 814
 - Gradient-Free, 803
 - Large-Scale, 810
 - Linear, 773
 - Methods
 - Search, 804
 - Nonlinear, 773
 - Nonsmooth, 814
 - Practical, 810
 - Search
 - Methods, 804
- Optimum
 - Global, 773
- Oral
 - Cavity, 48
- Ordinary
 - Discontinuity, 653
- Oren, 787
- Oren-Spedicato, 788
- Organic Pitch Range, 132
- Orthogonal
 - Set of Functions, 690, 712
- Orthogonal Matrices, 638
- Orthogonal Vectors, 638
- Orthogonality, 638
 - Function, 690
 - of Functions, 690
- Orthogonalization
 - Gram-Schmidt, 641
 - Modified, 641
 - Ordinary, 641
- Orthonormal
 - Basis, 396
 - Set of Functions, 691
 - Set of Vectors, 639
- Orthonormal Vectors, 638
- Orthonormality, 638
- Oscillatory
 - Discontinuity, 655
- Overfitting, 479
- Overpartitioning
 - k-Means, 364
- Overtraining, 479
- palatal
 - Sounds, 117
- Palate
 - Hard, 48
 - Soft, 44, 47
- Palato-Alveolar
 - Sounds, 117
- Papoulis, A., 85

- Papuan Language, 122
- Paralinguistic Pitch Range, 132
- PARCOR, 185
- PARCOR Coefficients, 176
- Parseval's
 - Theorem, 734
- Parseval's Theorem, 714
- Parsimony
 - Principle of, 350, 393
- Partan
 - Continuous, 802
 - Iterative Conjugate Gradient, 802
- Partial
 - Differentiation
 - Notation, 664
- Partial Correlation, *see* PARCOR
- Partial Correlation Coefficients, 176
- Partition, 208, 209
- Pathogenic Bacteria, 25
- Pattern
 - Recognition, 465
- PCA, 26, 394, 469
 - Kernel, *see* Kernel PCA
 - Nonlinear, 399
- PCM
 - μ -Law, 843
 - A-Law, 843
 - Adaptive Differential, 845
 - Linear, 842
 - Mu-Law, 843
 - PCMA, 843
 - PCMU, 843
 - Uniform, 842
- PCMA, 851
- PCMU, 851
- PCR, 24, 25
- Pearson, 474
 - Updates, 781
- Penalty Function, 318, 488
- Perception, 397
 - Auditory, 144
- Perceptual Linear Prediction, *see* PLP
- Perceptual Linear Predictive Analysis, 191
- Perikaryon, 51
- Periodic
 - Function, 663
 - Extension, 663
- Periodicity, 661
- Periodogram, 177, 735, 739
 - Estimate, 735
- Persian, *see* Languages
 - Empire, 129
- Phantoms, 537
- Pharyngeal
 - Sounds, 116
- Pharynx, 44, 47, 64
- Phon, 151
- Phonation, *see* Phone
 - Affricates, 110
 - Approximants, 110
 - Fricatives, 110
 - Lateral Resonant Contoids, 110
 - Nasal, 110
 - Nonsyllabic, 110
 - Normal, 110
 - Oral, 110
 - Resonants, 110
 - Rolls, 110
 - Sibilant, 110
 - Stops, 110
 - Syllabic, 110
 - Trills, 110
 - Vocoids, 110
- Phone, 76, 107, 108
 - Approximants, 124
 - Articulation, 108, 110
 - Place of, 110
 - Coordination, 108, 111
 - Diphthong, 124
 - Glide, 124
 - Glottal Stop, 124
 - Initiation, 108, 109
 - Liquid, 124
 - Offset, 161
 - Onset, 161
 - Phonation, 108
 - Unvoiced, 109
 - Voiced, 109
 - Standard Consonants, 124
 - Syllabic Consonants, 124
 - Triphthong, 124
- Phoneme, 107, 108
 - Affricate, 124
 - Diphthong, 124
 - Fricative, 124
 - Glide, 124
 - Liquid, 124
 - Semi-Vowel, 124
 - Whisper, 124
- Phonetic
 - Continuity, 138
 - Intonation, 132
 - Pitch, 132
 - Rate, 138
 - Rhythm, 138
 - Stress, 138
 - Tonality, 132
 - Tone, 132

- Vowels, 112
- Phonetics, 107
- Phonology, 107, 122, 129
- Phua, 787, 788
- Physiology, 397
- Piecewise
 - Continuity, 658
 - Smoothness, 658
- Pinna, 25
- Pitch, 50, 132, 144, 146, 149, 848
 - Cycle, 195
 - Linguistic Range, 132
 - Organic Range, 132
 - Paralinguistic Range, 132
- Pitch Scale, 169
- Pitch Variations, 138
- Place Theory, 145
- Plancherel's
 - Theorem, 734
- PLP, 157, 190, 191
 - Cepstra, 193
- Points
 - Stationary, 773
- Pointwise
 - Analyticity, 665
- Pointwise Convergence, *see* Convergence
- Poles
 - of Function, 686
- Polymerase Chain Reaction, *see* PCR
- POLYVAR, 620
- Pons, 54
- Positive
 - Definite, 787
- Positive Definite, 640
- Positive Definiteness, 640
- Post Alveolar
 - Sounds, 118
- Powell, 474, 783, 806, 808
- Powell's Method, 805
- Power
 - Spectral Density, 735, 739
 - Estimation, 735, 739
- Power Series, *see* Series
- Power Series Expansion, 683
- Praat, 628
- Practical
 - Issues
 - GMM, 451
 - HMM, 451
 - Neural Networks, 479
 - Sampling, 92
 - Errors, 92
- Pragmatics, 107
- pRAM, 465
- Pre-Emphasis, 153
 - Equal Loudness, 192, 193
- Pre-emphasis, 176
- Pre-Hilbert Space, 220
- Pre-Processing, 199, 561
- Primal
 - Problem, 835
- Primary Auditory Cortex, 66
- Principal Component Analysis, 394, *see* PCA
- Principle
 - Maximum
 - Entropy, 346
 - Minimum
 - Discriminability, 346
 - Parsimony, of, 350, 393
- Privacy, 624
- Probabilistic Random Access Memory, *see* pRAM
- Probability
 - Conditional, 224
 - Density Function, 229, 231
 - Joint, 233
 - Marginal, 233
 - Distribution, 247
 - Cumulative, 247
 - Cumulative Function, 235
 - Functions, 228
 - Integration, 227
 - Mass Function, 247
 - Measure, 221
 - Total, 225
- Probability Theory, 205
- Problem
 - Coin Toss, 414
 - Eigenvalue, 394
 - Generalized, 397
 - Sturm-Liouville, 85
- Problems
 - N -class, 518
 - Γ -class, 518
 - 2-class, 493
- Process
 - Hybridization, 24, 25
 - Markov, 268
 - Discrete, 268
- Processing
 - Image, 465
 - Signal, 143
- Product
 - Cartesian, 216
 - Complex
 - Variables, 649
 - Kronecker, 636
 - Space, 232

- Transformations, 238
- Product Space, 216
- Programming
 - Dynamic, 432
 - Integer, *see* Integer Programming
- Projected
 - Newton, 802
- Pronunciation
 - Vowels, 113
- Properties
 - Complex
 - Variables, 649
 - Definite
 - Integral, 673
 - Even
 - Function, 663
 - Integral
 - Definite, 673
 - Riemann, 673
 - Odd
 - Function, 663
 - Riemann
 - Integral, 673
- Property
 - Asymptotic
 - Equipartition, 274
- Prosencephalon, 54
- Prosodic
 - Features, 132
- Prosody, 107, 131, 132
- Protein, 52
- Protocol, 854
 - Transport
 - Real-time, *see* RTP
- PSD, 735, 739
 - Estimation, 735, 739
- Pseudo-Convex, 831
- Pseudo-Inverse, 640, 820
- PSTN, 617
- Psychophysical Power Law of Hearing, 193
- Pulmonic
 - Consonants, 115
- QCELP, 847
- Quadratic
 - Convergence, 792
 - Function, 773
 - Mean, 257
- Qualcomm, 847
- Quality
 - Audio, 627
 - Data, 627
 - Model, 534
- Quality Control
 - of Enrollment Utterances, 534
- Quantization, 155
 - Error
 - Amplitude, 85
- Quasi Newton, 474
- Quasi-Newton
 - BFGS, *see* BFGS
 - Condition, 779
 - Davidon
 - No Line Search, 790
 - DFP, *see* DFP
 - Hoshino, 789
 - Inexact
 - Line Search, 789
 - Initial Scaling, 788
 - Limited Memory, 811
 - Partially Separable, 811, 813
 - Pearson
 - Updates, 781
 - Sparse, 811, 812
- Quotient
 - Complex
 - Variables, 649
 - Set, 209
- Rényi
 - Entropy, 278, 279
- Rabiner, L.R., 124
- Radial Basis Function
 - Gaussian, *see* GRBF
 - Kernel, *see* Kernel
 - Neural Networks, *see* RBFN
 - RBF, 469
- Radon-Nikodým
 - Derivative, 230
 - Theorem, 229
- Random
 - Process
 - Bernoulli, 414
 - Variable
 - Bernoulli, 247
 - Combination, 234
 - Continuous, 227
 - Convergence, 250
 - Discrete, 247
 - i.i.d., 255, 256
- Random Variable
 - Continuous, 226
- Random Variables, 207
- Rank, 639
 - Full, 638, 820
- Ranvier Node, 53
- Ranvier Nodes, 53
- RASTA, 193, 571

- Rate
 - Equal Error, 589
 - Half Total Error, 590
 - Total Error, 590
- Ratio
 - Log-Likelihood, 288, 321
 - Minimum Tolerated, 396
- RBF, 469
- RBFN, 469
- Real Number
 - Extended, 215
- Real-time
 - Transport
 - Protocol, *see* RTP
- Recognition
 - Handwriting, 486
 - Online, 397
 - Image, 397, 486
 - Language, 411
 - Modalities, 12
 - Pattern, 465
 - Signature, 465
 - Speaker, *see* Speaker Recognition, 411, 479, 486
 - Speech, *see* Speech Recognition, 411, 465, 479, 486
- Recurrent
 - Neural Network, 477
 - Neural Networks, 476
- Redundancy, 274
- Reeves, 798
- Reflection Coefficients, 176, 186
- Region
 - Feasibility, *see* Feasibility Region
- Regression
 - Class, 606
- Regular
 - Function, 665
- Regularity
 - Assumption, 297, 826, 827, 831
 - Conditions, 296
- Relation
 - Completeness, 694
 - Equivalence, 208
 - Indiscernibility, 208
- Relative
 - Absolute Continuity of Measures, 229
 - Entropy, 286, 288, 291
 - Minimum, 346
 - Information, 321
 - Intensity, 150
- Relative Entropy, 273
- RelAtive SpecTrAl, *see* RASTA
- Relaxation
 - Stochastic, 367, 811
- Relevance Vector Machines, 553
- Representation
 - Results, 589
- Reproducing Kernel, *see* Kernel
- Residual
 - Information Criterion, *see* RIC
- Residues, 686
- Resolution, 155
- Results
 - Evaluation, 589
 - Identification, 593
 - Representation, 589
 - Verification, 589
- Retroflex
 - Sounds, 117
- Reverberation Compensation, 554
- Rhombencephalon, 54
- Rhythm, 138
- RIC, 349
- Riemann
 - Cauchy-Riemann
 - Conditions, 667
 - Theorem, 667–669
 - Integral, 673
 - Properties, 673
 - Lemma, 673
- Risk
 - Bayes, 488
 - Bounds, 493
 - Minimization, 488
 - Empirical, 492
 - Structural, 493
- RMS, 257
- RNN, 476, 477
- ROC Curve, 590
- Root Mean Square, 257
- Rosenbrock's Method, 804
- Rotokas, *see* Languages
- Rough
 - k-Means, 375
 - Set
 - Theory, 210
- RTP, 858
- RVM, 553
- Saddle Point, 773, 830
- SAFE, 849, 850
- Sample
 - Mean, 259
 - Space, 205
 - Variance, 253, 254, 256, 258
- Sample Space, 205
 - Atoms, 270, 282

- Sampler
 - Ideal, 98
- Sampling, 76, 152, 176
 - Audio, 77
 - Cyclic-Rate, 77
 - Multirate, 77
 - Periodic, 77
 - Pulse Width Modulated, 77
 - Random, 77
 - Theorem, 77
 - Convergence Criteria, 84
- Sampling Theorem, 79
 - Extended, 84
 - Extensions, 84
 - Generalized, 85
 - Whittaker-Kotelnikov-Shannon-Kramer, 85
 - WKS, 80
 - WKS, 85
- Sampling Theorem, The, 78
- Satellite Imaging, 486
- Scala, 145
- Scaling, 248
 - Variance
 - Discrete Random Variable, 249
- Scaling Associativity
 - Convolution, 689
- Scatter
 - Matrix, 402
- Schür Algorithm, 183
- Schür Recursion, 183
- Schmidt's Expansion Theorem, *see* Theorem
- Schwann Cell, 53
- Schwarz's Inequality, *see* Inequality
- Score
 - Fisher, 342
 - Statistic, 342
- Search
 - Methods, 804
- Second
 - Moment, 242
- Secondary Auditory Cortex, 68
- Segmentation, 3
 - Audio, 532
 - Speaker, 9, 549
- Selection
 - Model, 349
- Self Scaling Variable Metric, *see* SSVM
- Self-Dual, *see* Duality
- Self-Scaling Variable Metric, *see* SSVM
- Semantics, 107
- Semi-Supervised Learning, 390
- Separability
 - Linear, 493
- Sequence
 - Cauchy, 251
 - Classification, 328
 - Convergence, 250
 - Directional, 818, 828
 - Fundamental, 251
 - Limit, 651
- Sequential Interacting Multiple Models, 579
- Series
 - Fourier
 - convergence, 713
 - Laurent, 684
 - McLauren, 684
 - Power
 - Addition, 685
 - Division, 685
 - Multiplication, 685
 - Uniqueness, 685
 - Taylor, 683
 - Wavelet, 716
- Serotonin, 53
- Set, 205
 - Convex, 218
 - Fuzzy
 - Theory, 211
 - Quotient, 209
 - Rough
 - Theory, 210
 - Theory
 - Fuzzy, 211
 - Rough, 210
- Set Theory, 205, 207
- Sets
 - Borel, 214
- Settings, 131
- Sex
 - Classification, 8
- Shanno, 783, 787, 788
- Shannon, C.E., 80
- Shattering, 493
- Sheath, 53
- Sheep, 536
- Sherman-Morrison Inversion Formula, 642
- Shimmer, 8
- Short-Time
 - Gaussianization, 578
- Signal, 75
 - Enhancement, 199, 561
 - Non-Stationary, 76
 - Processing, 143
 - Stationary, 76
 - Time-Dependent, 75
- Signal Representation, 75
- Signature
 - Recognition, 465

- Verification, 486
- Silbo, 63, 137
- Silbo Gomero, 119
- Silence
 - Detection, 561
- SIMM, 579
- Simply Connected
 - Domain, 676
- Simulated
 - Annealing, 367, 474, 811
- Singular Value, 696
- Singular Value Decomposition, *see* SVD, 639
- Singular Values, 639
- Singularities
 - Isolated, 686
- SIVA, 620
- Skew, 245
- Skewness, 245
- Slack Variable, 500
- Smith, 806, 808
- Smooth
 - Functions, 773, 814
- Smoothness, 657
 - Piecewise, 658
- Soft Margin SVM, 501, 502
- Soft Palate, 44, 47
- Solution
 - Global, 474
- Solutions
 - Information Theory, 901
 - Integral Transforms, 904
 - Neural Networks, 902
- Sone, 151
- Sonority, 137
- Sound
 - Level, 564
- Sounds
 - Alveolar, 117
 - Plato-Alveolar, 117
 - Post Alveolar, 118
 - Bilabial, 118
 - Dental, 118
 - Epiglottal, 116
 - Glottal, 116
 - Labiodental, 118
 - palatal, 117
 - Pharyngeal, 116
 - Plato-Alveolar, 117
 - Post Alveolar, 118
 - Retroflex, 117
 - Uvular, 116
 - Velar, 116
 - Whisper, 119
 - Whistle, 119
- Source
 - Continuous, 269
 - Discrete, 267
 - Ergodic, 269
 - Information, 266
 - Markov, 268
 - Discrete, 268
 - Memoryless, 267, 413
 - Zero-Memory, 267
- Sources
 - Non-Unifilar, 268, 416, 420
 - Unifilar, 268, 416, 419
- Space
 - Banach, 218
 - Complete, 216, 220, 693
 - Countable, 207
 - Dot Product, 219, 220
 - Dual, *see* Dual Representation
 - Function, 236
 - Hilbert, 220
 - Infinite Dimensional, 219
 - Inner Product, 219, 220
 - Lebesgue, 236
 - Lebesgue Measurable, 216
 - Measurable, 209, 214
 - Metric, 217
 - Complete, 218
 - Multidimensional, 232
 - Normed Vector, 218
 - Pre-Hilbert, 220
 - Product, 216
 - Sample, 205
 - Atoms, 282
- Spanish Silbo, 119
- Speaker
 - Adaptation, 530
 - Over time, 601
 - Authentication, 3, 5
 - Biometrics, 3
 - Classification, 3, 8, 550
 - Detection, 11
 - Enrollment, 543
 - i-Vectors, *see* i-Vectors
 - Identification, 3, 7, 465, 548
 - Closed-Set, 7, 548
 - Open-Set, 7, 549
 - Independent Model, 529
 - Menagerie, 536
 - Model
 - Individual, 526
 - Modeling, 525
 - Recognition, 3, 411, 479, 486, 543
 - Branches, 5, 543
 - History, 3

- Knowledge-Based, 15, 548
- Manifestations, 5, 543
- Modalities, 12
- Text-Dependent, 12, 546
- Text-Independent, 13
- Text-Prompted, 14, 546
- Segmentation, 3, 9, 549
- Target, 5, 313, 314
- Test, 5, 313, 314
- Tracking, 3, 11
- Verification, 3, 5, 506, 508, 544
- Speaker Diarization, 19
- Speaker Factors, 532
- Speaker Indexing, 19
- Speaker Model Synthesis, 578
- Speaker Normalization, 581, 583
- Speaker Recognition, 69
 - Branches
 - Compound, 5
 - Simple, 5
- Speaker Space, 532
- Specific Acoustic Impedance, 149, 150
- Spectral
 - Filtering, 571
 - Flatness, 168
 - Representation, 79
- Spectral Analysis, 157, 168, 191
- Spectral Envelope, 8
- Spectrogram, 87
 - Narrowband, 88
 - Wideband, 88
- Spedicato, 788
- Speech
 - Anatomy, 43
 - Co-Articulation, 140
 - Detection, 561
 - Features
 - Temporal, 140
 - Header
 - Resources, *see* SPHERE
 - Perception, 69
 - Recognition, 3, 411, 465, 479, 486
 - Signal Representation, 75
 - Synthesis, 411
 - Temporal Features, 140
 - Waveform, 87
- Speech Production, 43
- Speex, 847
- SPHERE, 849, 850
- SPIDRE, 618
- Spoofting, 4, 13, 547, 625
- Square-root Filtering, *see* Cholesky
 - Factorization
- SRAPI, 855
- ssDNA, *see* DNA Single Strand
- SSVM, 474, 787–789
- Stability
 - Numerical, 623, 813
- Standard
 - Audio
 - Encapsulation Formats, 849
 - Format Encapsulation, *see* SAFE
 - Standard Intensity Threshold, 150, 172
- State
 - Confluent, 458
- State-Space, 479
- States, 421, 774
- Stationary
 - Points, 773
- Stationary Signal, 76
- Statistic, 251
 - Consistent, 253
 - Efficiency of, 252
 - Efficient, 252
 - Score, 342
 - Sufficient, 252
- Statistical
 - Independence, 226
 - Mutual, 226
 - Moment
 - Estimation, 253
 - First, 248
- Statistical Moments, 239
- Statistics
 - High Order, 381
 - Sufficient, 251
- Steifel, 798
- Steinberg, J.C., 146
- Stevens, S.S., 146
- Stochastic
 - Matrix, 643
 - Relaxation, 367, 811
- Streaming, 852
- Stress, 138
 - Linguistic, 135
- Strict
 - Minimizer
 - Local, 774
- Strictly
 - Concave
 - Function, 661
 - Convex
 - Function, 658, 661
- Stricture
 - Degree of, 110
- Stride, 33
- Strong
 - Completeness, 218

- Convergence, 251
- Strong Law of Large Numbers, 255, 256
- Structural Risk Minimization, 493
- Sturm-Liouville Problem, 85, 709, 712
- Subgradients, 814
- Subset, 205
 - Convex, 218
 - Distance, 217
- Subsets
 - Lebesgue Measurable, 215
- Sufficient
 - Statistics, 251
- Sulci, 54
- Sulcus, 54
- Superset, 205
- Supervised
 - Clustering, 357
- Support Vector Machines, *see* SVM
- Support Vectors, 495, 823
 - Definition, 485
- Suprasegmental, 131
- Suprasegmental Features, 107
- SVAPI, 855
- SVD, 579, 639
- SVM, 348, 479, 485, 503, 583, 695, 823
 - L_1 Soft Margin, 501
 - L_2 Soft Margin, 502
 - Definition, 485
- Switchboard I, 618
- Switchboard II, 618
- Syllable
 - Coda, 133
 - Nucleus, 133
 - Onset, 133
- Sylvian Fissure, 66
- Symmetric
 - Channel
 - Binary, 284
 - Kernel, *see* Kernel
 - Degenerate, *see* Kernel
- Synapse, 51
- Syntax, 107, 129
- Synthesis
 - Speech, 411
- Systems
 - Auditory, 49
 - Ear, 49
 - Auditory Cortex, 66
 - Large-Scale, 628
 - Speech Production, 43
 - Vocal, 44
- T-NETIX, 855
- Tactical Speaker Identification, 619
- Target
 - Speaker, 5, 313, 314
- Taylor
 - Series, 683, 777, 817, 828
- TDNN, 465, 477, 479
- Teeth, 48
- Teleconferencing, 21
- Telencephalon, 54
- Tempo, 140
- Temporal Features, 140
- Tensor, 472
- TER, 590
- Terminal Buttons, 53
- Tertiary Auditory Cortex, 68
- Test
 - Speaker, 5, 313, 314
- Testing
 - Hypothesis, 313
- Texas Instruments, 855
- Text-Dependent, 12, 546
- Text-Independent, 13
 - Speaker Identification, 479
- Text-Prompted, 14, 546
- Thalamus, 54
- Theorem
 - Bayes, 225
 - Cauchy
 - Integral, 676, 683
 - Cauchy-Riemann, 667–669
 - de Moivre, 650
 - Hilbert's Expansion, 504, 698
 - Mean Value, 669, 670, 676
 - Mercer's Expansion, 504, 698
 - Morera, 683
 - Parseval, 714
 - Fourier Series, 714
 - Radon-Nikodým, 229
 - Sampling, 79
 - Extended, 84
 - Extensions, 84
 - Sampling, The, 78
 - Schmidt's Expansion, 504, 698
 - Whittaker-Kotelnikov-Shannon, 80
 - WKS, 80
- Theory
 - Information, 265
 - Place, 145
 - Set
 - Fuzzy, 211
 - Rough, 210
 - VC, *see* VC Theory
- Third
 - Moments, 245
- Thomas Bayes, 225

- Thymine, *see* DNA Thymine
- Timbre, 144, 151
- Time Lapse Effects, 595
- Time-Delay
 - Neural Network, 477
- Time-Delay Neural Network, *see* TDNN
- Time-Dependent Signal, 75
- TIMIT, 617
- Toeplitz Matrix, 183
- Tonality, 132
- Tone
 - Removal, 579, 580
- Tongue, 44, 64
- Tools
 - HTK, 628
 - Praat, 628
 - Voicebox Toolkit, 628
- Total
 - Factors, 532
 - Variability, 532
 - Space, 532
- Total Error Rate, 590
- Total Probability, 225
- Trachea, 44
- Tracking, 3
 - Speaker, 11
- Tractability
 - Model, 449
- Training
 - GMM, 451
 - HMM, 423, 451
 - Neural Network, 473
- Transform
 - Complex Fourier, 722
 - Complex Short-Time Fourier, 740
 - Discrete Cosine, 748
 - Discrete Fourier
 - Inverse, 732
 - Discrete Short-Time Fourier, 746
 - Discrete-Term Short-Time Fourier, 744
 - Discrete-Time Short-Time Fourier, 744
 - Fourier, 580, 717, 722
 - Complex, 80
 - Cosine Discrete, 748
 - Discrete, 731
 - Discrete Cosine, 748
 - Short-Term, 740
 - Short-Term Discrete, 746
 - Short-Term: Discrete-Time, 744
 - Short-Time, 740
 - Short-Time Discrete, 746
 - Short-Time: Discrete-Time, 744
 - Fourier Integral, 722
 - Gabor, 167, 740
 - Integral, 647, 717
 - Fourier, 722
 - General, 695
 - Laplace, 717
 - Inverse
 - Laplace, 720
 - Laplace, 717
 - Inversion, 720
 - z, 717, 750
- Transformation, 238
 - Inverse Image, 238
 - Karhunen Loève, *see* KLT
 - Karhunen-Loève, 394
 - Linear, 403
 - Measurable, 238
 - One-to-one, 238
 - Orthogonal
 - Linear, 394
 - Product of, 238
- Transformations, 238, 393
- Translation, 248
- Transpose, 635
- Tree
 - k-d, 368, 375
 - k-dimensional, *see* k-d tree
- Trees
 - Decision, 331
- Trellis Diagram, 428
- Tremolo, 151
- Triangular
 - Inequality
 - Complex Plane, 649
- Triangular Inequality, *see* Inequality
- Triangular Window, 165
- Trick
 - Kernel, 506
- Trigonometric
 - Function
 - Analyticity, 672
- Triphosphate, *see* DNA Triphosphate
- Truncation
 - Error, 102
- TSID, 619
- Twiddle
 - Factor, 733
- Two-Wire Telephone Channel, 532
- Tympanic Membrane, 49
- UBM, 6
- Uncertainty, 269
 - Continuous Sources, 284
 - Discrete Sources, 269, 270
- uncompressed, 851
- Underflow, 414

- Understanding
 - Language, 411
- Unifilar, 268, 416
 - Source, 419
- Uniform
 - Density Function, 285
 - PCM, 842
- Uniform Convergence, *see* Convergence
- Unit
 - Likelihood, 453
- Unitary Matrices, 638
- Universal
 - Background
 - Model, 527
- Universal Background Model, 6
- Unsupervised
 - Clustering, 341, 357, 359
- Updates
 - Pearson, 781
- Uvula, 44
- Uvular
 - Sounds, 116
- VAD, 561
- Validation
 - Cross, *see* Cross-Validation
- van Essen, D., 55
- Vapnik, Vladimir, 486
- Vapnik-Chervonenkis
 - Dimension, *see* VC Dimension
 - Theory, *see* VC Theory
- Variable
 - Bit Rate, 852
 - Continuous
 - Limit: Function, 651
- Variable Metric, 474
 - Methods, 779
 - Self Scaling, *see* SSVM
 - Self-Scaling, *see* SSVM
- Variables
 - Dual, 816
- Variance, 242
 - Estimation, 258
 - Random Variable
 - Discrete, 249
 - Sample, 258
- Variance Matrix, 260
- Variance-Covariance Matrix, 260
- Variations
 - Illumination, 26
- VC
 - Dimension, 350, 485, 486, 493
 - Theory, 493
- Vector
 - Euclidean Norm, 636
 - Norm, 636
 - Normal, *see* Normal Vector
 - Representation, 642
- Vector Quantization, 358
- Vectors
 - Conjugate, 638
 - Orthogonal, 638
 - Orthonormal, 638, 639
- Velar
 - Sounds, 116
- Velum, 44
- Venn Diagram, 205
- Ventriculus Tertius, 54
- Verification, 3, 5, 544
 - Signature, 486
- Verification Results, 589
- Vestibulocochlear Nerve Bundle, 49
- Vibrato, 151
- Video Indexing, 19
- Viterbi, 432
- Vocal Chords, 44
- Vocal Folds, 44
- Vocal Source Features, 195
- Vocal System, 43
- Vocal Tract, 64
 - Length, 8
 - Normalization, 573
 - Shape, 8
- Vocal Tracy, 44
- Voice
 - Activity Detection, 561
 - Creaky, *see* Laryngealization
 - Recognition, 3
- Voicebox Toolkit, 628
- Voiced
 - Implosives, 120
- Volume, 564
- Von Neumann Computer, 465
- Vorbis
 - OGG, 844
- Vowel, 8, 119, 127
- Vowels, 112
 - Pronunciation, 113
- VQ, 358
- VTLN, 573
- Warping
 - Bark Frequency, 191
 - Feature, 576
 - Frequency, 169
 - Magnitude, 172
- WAV, 849
 - Audio

- Encapsulation, 849
- Waveform
 - Speech, 87
- Wavelet
 - Biorthogonal, 194
 - Coiflets, 194
 - Daubechies, 194
 - DMeyer, 194
 - Haar, 194
 - Octave Coefficients Of Residues, 195
 - Reverse Biorthogonal, 194
 - Series, 716
 - Symlets, 194
- WCCN, 583
- Weak
 - Completeness, 218
 - Convergence, 250
- Weak Law of Large Numbers, 255
- Weighted Euclidean Distance, 303
- Welch Window, 163
- Wernicke's Area, 60, *see* Brain
- Whisper, 109, 119
- whisper, 125
- Whistle, 63, 119, 137
- White Matter, 54
- Whittaker, E.T., 80
- Whittaker, J.M., 80
- Wideband Spectrogram, 88
- Wiener Filter, 580
- Window
 - Bartlett, 163, 165
 - Blackman, 165
 - Gauss, 167
 - Hamming, 162, 580
 - Hann, 163
 - Low-pass, 176
 - Triangular, 165
 - Welch, 163
- Windowing, 161
- Within Class Covariance Normalization, 583
- WKS
 - Sampling Theorem, 80
- WKSK
 - Sampling Theorem, 80
- WMAP, 583
- WMLLR
 - Adaptation, 607
- WOCOR, 195
- Wolfe, 783
 - Dual Representation, *see* Duality
 - Duality Theorem, 831
- Wolfe's Duality, *see* Duality
- Wolfe-Powell
 - Conditions, 783
- Wolves, 536
- Wood
 - Direct Search, 804
- World MAP, 583
- World Maximum A-Posteriori, 583
- World Wide Web, *see* WWW
- Worms, 537
- Wrappers
 - k-Means, 368
- WWW, 509
- x-Means
 - Modified, 372
- XOR, 481
- YOHO, 618
- Yule-Walker Equations, 183
- z-Transform, 750
- Zero Entropy, 272
- Zero Normalization, 581
- Zero-Memory, *see* Memoryless
- Zero-Memory Source, 267
- Zeros
 - of a Function, 685
- Zoutendijk, 802