CONTENTS

PREFACE ix

I MOTIVATION AND THE BASICS

1 INTRODUCTION 3
   1.1 Least Squares and Regularized Regression / 4
   1.2 Lasso: Survival of the Bigger / 6
   1.3 Thresholding the Sample Covariance Matrix / 9
   1.4 Sparse PCA and Regression / 10
   1.5 Graphical Models: Nodewise Regression / 13
   1.6 Cholesky Decomposition and Regression / 13
   1.7 The Bigger Picture: Latent Factor Models / 15
   1.8 Further Reading / 17

2 DATA, SPARSITY, AND REGULARIZATION 21
   2.1 Data Matrix: Examples / 22
   2.2 Shrinking the Sample Covariance Matrix / 26
   2.3 Distribution of the Sample Eigenvalues / 29
   2.4 Regularizing Covariances Like a Mean / 30
   2.5 The Lasso Regression / 32
   2.6 Lasso: Variable Selection and Prediction / 36
CONTENTS

2.7 Lasso: Degrees of Freedom and BIC / 37
2.8 Some Alternatives to the Lasso Penalty / 38

3 COVARIANCE MATRICES

3.1 Definition and Basic Properties / 45
3.2 The Spectral Decomposition / 49
3.3 Structured Covariance Matrices / 53
3.4 Functions of a Covariance Matrix / 56
3.5 PCA: The Maximum Variance Property / 61
3.6 Modified Cholesky Decomposition / 63
3.7 Latent Factor Models / 67
3.8 GLM for Covariance Matrices / 73
3.9 GLM via the Cholesky Decomposition / 76
3.10 GLM for Incomplete Longitudinal Data / 79
   3.10.1 The Incoherency Problem in Incomplete Longitudinal Data / 79
   3.10.2 The Incomplete Data and The EM Algorithm / 81
3.11 A Data Example: Fruit Fly Mortality Rate / 84
3.12 Simulating Random Correlation Matrices / 89
3.13 Bayesian Analysis of Covariance Matrices / 91

II COVARIANCE ESTIMATION: REGULARIZATION

4 REGULARIZING THE EIGENSTRUCTURE

4.1 Shrinking the Eigenvalues / 100
4.2 Regularizing The Eigenvectors / 105
4.3 A Duality between PCA and SVD / 107
4.4 Implementing Sparse PCA: A Data Example / 110
4.5 Sparse Singular Value Decomposition (SSVD) / 112
4.6 Consistency of PCA / 114
4.7 Principal Subspace Estimation / 118
4.8 Further Reading / 119

5 SPARSE GAUSSIAN GRAPHICAL MODELS

5.1 Covariance Selection Models: Two Examples / 122
5.2 Regression Interpretation of Entries of $\Sigma^{-1}$ / 124
5.3 Penalized Likelihood and Graphical Lasso / 126
CONTENTS

5.4 Penalized Quasi-Likelihood Formulation / 131
5.5 Penalizing the Cholesky Factor / 132
5.6 Consistency and Sparsistency / 136
5.7 Joint Graphical Models / 137
5.8 Further Reading / 139

6 BANDING, TAPERING, AND THRESHOLDING 141
6.1 Banding the Sample Covariance Matrix / 142
6.2 Tapering the Sample Covariance Matrix / 144
6.3 Thresholding the Sample Covariance Matrix / 145
6.4 Low-Rank Plus Sparse Covariance Matrices / 149
6.5 Further Reading / 150

7 MULTIVARIATE REGRESSION: ACCOUNTING FOR CORRELATION 153
7.1 Multivariate Regression and LS Estimators / 154
7.2 Reduced Rank Regressions (RRR) / 156
7.3 Regularized Estimation of B / 158
7.4 Joint Regularization of (B, \Omega) / 160
7.5 Implementing MRCE: Data Examples / 163
7.5.1 Intraday Electricity Prices / 163
7.5.2 Predicting Asset Returns / 165
7.6 Further Reading / 167

BIBLIOGRAPHY 171
INDEX 181