

INVERSE PROBLEMS IN GEOPHYSICS  
GEOS 567

**A Set of Lecture Notes**

**by**

**Professors Randall M. Richardson and George Zandt  
Department of Geosciences  
University of Arizona  
Tucson, Arizona 85721**

**Revised and Updated Fall 2009**

## TABLE OF CONTENTS

PREFACE .....		v
CHAPTER 1: INTRODUCTION .....		1
1.1	Inverse Theory: What It Is and What It Does .....	1
1.2	Useful Definitions .....	2
1.3	Possible Goals of an Inverse Analysis .....	3
1.4	Nomenclature .....	4
1.5	Examples of Forward Problems .....	7
1.5.1	Example 1: Fitting a Straight Line .....	7
1.5.2	Example 2: Fitting a Parabola .....	8
1.5.3	Example 3: Acoustic Tomography .....	9
1.5.4	Example 4: Seismic Tomography .....	10
1.5.5	Example 5: Convolution .....	10
1.6	Final Comments .....	11
CHAPTER 2: REVIEW OF LINEAR ALGEBRA AND STATISTICS .....		12
2.1	Introduction .....	12
2.2	Matrices and Linear Transformations .....	12
2.2.1	Review of Matrix Manipulations .....	12
2.2.2	Matrix Transformations .....	15
2.2.3	Matrices and Vector Spaces .....	19
2.2	Probability and Statistics .....	21
2.3.1	Introduction .....	21
2.3.2	Definitions, Part 1 .....	21
2.3.3	Some Comments on Applications to Inverse Theory .....	24
2.3.4	Definitions, Part 2 .....	25
CHAPTER 3: INVERSE METHODS BASED ON LENGTH .....		31
3.1	Introduction .....	31
3.2	Data Error and Model Parameter Vectors .....	31
3.3	Measures of Length.....	31
3.4	Minimizing the Misfit: Least Squares.....	33
3.4.1	Least Squares Problem for a Straight Line .....	33
3.4.2	Derivation of the General Least Squares Solution .....	36
3.4.3	Two Examples of Least Squares Problems .....	38
3.4.4	Four-Parameter Tomography Problem .....	40
3.5	Determinacy of Least Squares Problems .....	42
3.5.1	Introduction .....	42
3.5.2	Even-Determined Problems: $M = N$ .....	43
3.5.3	Overdetermined Problems: Typically, $N > M$ .....	43
3.5.4	Underdetermined Problems: Typically $M > N$ .....	43
3.6	Minimum Length Solution.....	44
3.6.1	Background Information .....	44
3.6.2	Lagrange Multipliers .....	45
3.6.3	Application to the Purely Underdetermined Problem .....	48

3.6.4	Comparison of Least Squares and Minimum Length Solutions .....	50
3.6.5	Example of Minimum Length Problem .....	50
3.7	Weighted Measures of Length .....	51
3.7.1	Introduction .....	51
3.7.2	Weighted Least Squares.....	52
3.7.3	Weighted Minimum Length .....	55
3.7.4	Weighted Damped Least Squares .....	57
3.8	A Priori Information and Constraints .....	58
3.8.1	Introduction .....	58
3.8.2	A First Approach to Including Constraints .....	59
3.8.3	A Second Approach to Including Constraints .....	61
3.8.4	Example From Seismic Receiver Functions .....	64
3.9	Variance of the Model Parameters.....	65
3.9.1	Introduction .....	65
3.9.2	Application to Least Squares .....	65
3.9.3	Application to the Minimum Length Problem .....	66
3.9.4	Geometrical Interpretation of Variance .....	66
CHAPTER 4:	LINEARIZATION OF NONLINEAR PROBLEMS.....	70
4.1	Introduction .....	70
4.2	Linearization of Nonlinear Problems .....	70
4.3	General Procedure for Nonlinear Problems .....	73
4.4	Three Examples .....	73
4.4.1	A Linear Example .....	73
4.4.2	A Nonlinear Example .....	75
4.4.3	Nonlinear Straight-Line Example .....	81
4.5	Creeping vs Jumping ( <i>Shaw and Orcutt, 1985</i> ) .....	86
CHAPTER 5:	THE EIGENVALUE PROBLEM .....	89
5.1	Introduction .....	89
5.2	The Eigenvalue Problem for Square ( $M \times M$ ) Matrix $\mathbf{A}$ .....	89
5.2.1	Background .....	89
5.2.2	How Many Eigenvalues, Eigenvectors? .....	90
5.2.3	The Eigenvalue Problem in Matrix Notation .....	92
5.2.4	Summarizing the Eigenvalue Problem for $\mathbf{A}$ .....	94
5.3	Geometrical Interpretation of the Eigenvalue Problem for Symmetric $\mathbf{A}$ .....	95
5.3.1	Introduction .....	95
5.3.2	Geometrical Interpretation .....	96
5.3.3	Coordinate System Rotation .....	100
5.3.4	Summarizing Points .....	101
5.4	Decomposition Theorem for Square $\mathbf{A}$ .....	102
5.4.1	The Eigenvalue Problem for $\mathbf{A}^T$ .....	102
5.4.2	Eigenvectors for $\mathbf{A}^T$ .....	103
5.4.3	Decomposition Theorem for Square Matrices .....	103
5.4.4	Finding the Inverse $\mathbf{A}^{-1}$ for the $M \times M$ Matrix $\mathbf{A}$ .....	110
5.4.5	What Happens When There Are Zero Eigenvalues? .....	111
5.4.6	Some Notes on the Properties of $\mathbf{S}_P$ and $\mathbf{R}_P$ .....	114
5.5	Eigenvector Structure of $\mathbf{m}_{LS}$ .....	115
5.5.1	Square Symmetric $\mathbf{A}$ Matrix With Nonzero Eigenvalues .....	115
5.5.2	The Case of Zero Eigenvalues .....	117
5.5.3	Simple Tomography Problem Revisited .....	118

CHAPTER 6:	SINGULAR-VALUE DECOMPOSITION (SVD)	123
6.1	Introduction	123
6.2	Formation of a New Matrix <b>B</b>	123
6.2.1	Formulating the Eigenvalue Problem With <b>G</b>	123
6.2.2	The Role of $\mathbf{G}^T$ as an Operator	124
6.3	The Eigenvalue Problem for <b>B</b>	125
6.3.1	Properties of <b>B</b>	125
6.3.2	Partitioning <b>W</b>	126
6.4	Solving the Shifted Eigenvalue Problem	127
6.4.1	The Eigenvalue Problem for $\mathbf{G}^T\mathbf{G}$	127
6.4.2	The Eigenvalue Problem for $\mathbf{G}\mathbf{G}^T$	128
6.5	How Many $\eta_i$ Are There, Anyway??	129
6.5.1	Introducing $P$ , the Number of Nonzero Pairs $(+\eta_i, -\eta_i)$	130
6.5.2	Finding the Eigenvector Associated with $-\eta_i$	131
6.5.3	No New Information From the $-\eta_i$ System	131
6.5.4	What About the Zero Eigenvalues $\eta_i$ 's, $i = 2(P + 1), \dots, N + M$ ?	132
6.5.5	How Big is $P$ ?	133
6.6	Introducing Singular Values	134
6.6.1	Introduction	134
6.6.2	Definition of the Singular Value	135
6.6.3	Definition of $\Lambda$ , the Singular-Value Matrix	135
6.7	Derivation of the Fundamental Decomposition Theorem for General <b>G</b> ( $N \times M, N \neq M$ )	137
6.8	Singular-Value Decomposition (SVD)	138
6.8.1	Derivation of Singular-Value Decomposition	138
6.8.2	Rewriting the Shifted Eigenvalue Problem	140
6.8.3	Summarizing SVD	141
6.9	Mechanics of Singular-Value Decomposition	142
6.10	Implications of Singular-Value Decomposition	143
6.10.1	Relationships Between <b>U</b> , <b>U<sub>P</sub></b> , and <b>U<sub>0</sub></b>	143
6.10.2	Relationships Between <b>V</b> , <b>V<sub>P</sub></b> , and <b>V<sub>0</sub></b>	144
6.10.3	Graphic Representation of <b>U</b> , <b>U<sub>P</sub></b> , <b>U<sub>0</sub></b> , <b>V</b> , <b>V<sub>P</sub></b> , and <b>V<sub>0</sub></b> Spaces	145
6.11	Classification of $\mathbf{d} = \mathbf{G}\mathbf{m}$ Based on $P$ , $M$ , and $N$	146
6.11.1	Introduction	146
6.11.2	Class I: $P = M = N$	147
6.11.3	Class II: $P = M < N$	147
6.11.4	Class III: $P = N < M$	148
6.11.5	Class IV: $P < \min(N, M)$	149
CHAPTER 7:	THE GENERALIZED INVERSE AND MEASURES OF QUALITY	150
7.1	Introduction	150
7.2	The Generalized Inverse Operator $\mathbf{G}_g^{-1}$	152
7.2.1	Background Information	152
7.2.2	Class I: $P = N = M$	152
7.2.3	Class II: $P = M < N$	153
7.2.4	Class III: $P = N < M$	161
7.2.5	Class IV: $P < \min(N, M)$	164
7.3	Measures of Quality for the Generalized Inverse	166
7.3.1	Introduction	166
7.3.2	The Model Resolution Matrix <b>R</b>	167
7.3.3	The Data Resolution Matrix <b>N</b>	171

7.3.4	The Unit (Model) Covariance Matrix $[\text{cov}_u \mathbf{m}]$ .....	176
7.3.5	A Closer Look at Stability .....	176
7.3.6	Combining $\mathbf{R}$ , $\mathbf{N}$ , $[\text{cov}_u \mathbf{m}]$ .....	180
7.3.7	An Illustrative Example .....	181
7.4	Quantifying the Quality of $\mathbf{R}$ , $\mathbf{N}$ , and $[\text{cov}_u \mathbf{m}]$ .....	184
7.4.1	Introduction .....	184
7.4.2	Classes of Problems .....	184
7.4.3	Effect of the Generalized Inverse Operator $\mathbf{G}_g^{-1}$ .....	185
7.5	Resolution Versus Stability .....	187
7.5.1	Introduction .....	187
7.5.2	$\mathbf{R}$ , $\mathbf{N}$ , and $[\text{cov}_u \mathbf{m}]$ for Nonlinear Problems .....	189
CHAPTER 8: VARIATIONS OF THE GENERALIZED INVERSE .....		195
8.1	Linear Transformations .....	195
8.1.1	Analysis of the Generalized Inverse Operator $\mathbf{G}_g^{-1}$ .....	195
8.1.2	$\mathbf{G}_g^{-1}$ Operating on a Data Vector $\mathbf{d}$ .....	197
8.1.3	Mapping Between Model and Data Space: An Example .....	198
8.2	Including Prior Information, or the Weighted Generalized Inverse .....	200
8.2.1	Mathematical Background .....	200
8.2.2	Coordinate System Transformation of Data and Model Parameter Vectors .....	203
8.2.3	The Maximum Likelihood Inverse Operator, Resolution, and Model Covariance .....	204
8.2.4	Effect on Model- and Data-Space Eigenvectors .....	208
8.2.5	An Example .....	210
8.3	Damped Least Squares and the Stochastic Inverse .....	216
8.3.1	Introduction .....	216
8.3.2	The Stochastic Inverse .....	216
8.3.3	Damped Least Squares .....	220
8.4	Ridge Regression .....	225
8.4.1	Mathematical Background .....	225
8.4.2	The Ridge Regression Operator .....	226
8.4.3	An Example of Ridge Regression Analysis .....	228
8.5	Maximum Likelihood .....	232
8.5.1	Background .....	232
8.5.2	The General Case .....	235
CHAPTER 9: CONTINUOUS INVERSE THEORY AND OTHER APPROACHES .....		239
9.1	Introduction .....	239
9.2	The Backus–Gilbert Approach .....	240
9.3	Neural Networks .....	248
9.4	The Radon Transform and Tomography (Approach 1) .....	251
9.4.1	Introduction .....	251
9.4.2	Interpretation of Tomography Using the Radon Transform .....	254
9.4.3	Slant-Stacking as a Radon Transform (following <i>Claerbout</i> , 1985) .....	255
9.5	A Review of the Radon Transform (Approach 2) .....	259
9.6	Alternative Approach to Tomography .....	262

## PREFACE

This set of lecture notes has its origin in a nearly incomprehensible course in inverse theory that I took as a first-semester graduate student at MIT. My goal, as a teacher and in these notes, is to present inverse theory in such a way that it is not only comprehensible but useful.

Inverse theory, loosely defined, is the fine art of inferring as much as possible about a problem from all available information. Information takes both the traditional form of data, as well as the relationship between actual and predicted data. In a nuts-and-bolt definition, it is one (some would argue the best!) way to find and assess the quality of a solution to some (mathematical) problem of interest.

Inverse theory has two main branches dealing with *discrete* and *continuous* problems, respectively. This text concentrates on the discrete case, covering enough material for a single-semester course. A background in linear algebra, probability and statistics, and computer programming will make the material much more accessible. Review material is provided on the first two topics in Chapter 2.

This text could stand alone. However, it was written to complement and extend the material covered in the supplemental text for the course, which deals more completely with some areas. Furthermore, these notes make numerous references to sections in the supplemental text. Besides, the supplemental text is, by far, the best textbook on the subject and should be a part of the library of anyone interested in inverse theory. The supplemental text is:

**Geophysical Data Analysis: Discrete Inverse Theory (Revised Edition)**  
by William Menke, Academic Press, 1989.

The course format is largely lecture. We may, from time to time, read articles from the literature and work in a seminar format. I will try to schedule a couple of guest lectures in applications. Be forewarned. There is a lot of homework for this course. They are occasionally very time consuming. I make every effort to avoid pure algebraic nightmares, but my general philosophy is summarized below:

*I hear, and I forget.*  
*I see, and I remember.*  
*I do, and I understand.*  
– Chinese Proverb

I try to have you do a “simple” problem by hand before turning you loose on the computer, where all realistic problems must be solved. You will also have access to existing code and a computer account on a SPARC workstation. You may use and modify the code for some of the homework and for the term project. The term project is an essential part of the learning process and, I hope, will help you tie the course work together. Grading for this course will be as follows:

60% Homework  
30% Term Project  
10% Class Participation

Good luck, and may you find the trade-off between stability and resolution less traumatic than most, on average.

Randy Richardson  
August 2009