IMPROVED LINEAR REGRESSION WITH PROCESS APPLICATIONS

Dissertation by
Tyler Reed Holcomb

In Partial Fulfillment of the Requirements
for the Degree of
Doctor of Philosophy

California Institute of Technology
Division of Chemistry & Chemical Engineering
Pasadena, California

1993
(Defended 3 December 1992)

"A Ph.D. does not mean you have mastered the fundamentals of a field, created new knowledge, or furthered human understanding; these should be your goals when working on a Ph.D. A Ph.D. means you can choose to do something difficult and keep at it for five years."

— Dr. Terry Holcomb, in the presence of an apparently disinterested adolescent.

Thanks Dad, I really was listening.

Acknowledgments

I am deeply indebted to my co-workers Paul Schlosser, Hector Budman, and Chris Webb for their years of patient toil: Paul with the "little beasties," Hector and Chris with the "d**n reactor." The physics graduate students have a saying: experimentalists get their hands dirty, theorists get their minds dirty. Without their handiwork, my mind could never have gotten so filthy. I thank Paul for helping me write section 11.2 and for allowing me to copy appendix D in whole cloth. I thank Hector for his contributions to chapter 10. I gratefully acknowledge the profound contributions of my research advisor Manfred Morari. When a prior research effort came to a dead halt, he had the forthrightness to ask, "Well, what are you good at?" This thesis is a partial response to his question. His insistence on rigor and his belief that rigor should be used for practical applications, not just for its own sake, have taught me much.

This work would not exist were it not for the unselfish love of my parents. During many a dark day, a letter or e-mail from them was the bright spot; during the difficult times their support was unwavering. These few words can not express the depth of my gratitude. I also want to thank three wonderful teachers from Denton, Texas: Coach Ken Garland for showing me my inner strength, Ms. Charlotte Scroggs for her encouragement and hours of volunteer efforts, and Mrs. W. Burlage for never compromising her standards.

I have made so many friends at Caltech that I do not have room to name them all. Kudos to all the Franciscans and the Rude Brothers for never allowing life to be boring. I specifically want to thank Barb Wyslouzil for her love and understanding, and Achim Ditzen and Frank Doyle for being close friends and wonderful roomates. Growing-up in a small town, I was inculcated with the sadly-out-of-fashion idea "if

you're not part of the solution, you're part of the problem." I want to thank Lynn Hildeman for encouraging me to be active on the Graduate Student Council, the Graduate Office (Arden Albee and Carol Mastin in particular) for working so hard on solutions, and all the wonderful GSC directors and other student volunteers who have helped me make my small contributions to graduate student life. Foremost I thank Alan Blanchard and Stein Sigurdsson for preparing me for two difficult but vital years as Chair of the Graduate Review Board. Holding the scales of justice can be a terrifyingly awesome task; I wish Michael Malak well as he bravely shoulders this burden.

My research has not occurred in isolation; I have cherished my interactions with Amir, Tony, Pete, Upi, Ramin, Richard C., John, Lionel, Ravinder, Jay, Maurice, Marc, Richard B., Thanos, Frank A., Yasushi, George, Nikos, Alex, Simone, Mayuresh, Cris, Rick, Vasily, and all my wonderful other co-workers. I specifically want to thank Håkan Hjalmarsson and his probabilistic hotline ("where you might get an answer") for help with statistical distributions and essential bits of probablistic arcana. I thank the National Science Foundation for three years of support when the research results were fewer and further between and the Caltech Consortium in Chemistry and Chemical Engineering for supporting the crucial early phases of this research.

Lastly, I thank my loving wife, Michal Peri, for her support (and food) during this effort. Although she entered my life after this work had "turned the corner," I look forward to a lifetime of turning corners with her.

Ahni l'dodi, v'dodi li.

IMPROVED BIASED REGRESSION WITH PROCESS APPLICATIONS

Tyler Reed Holcomb

Abstract

This work examines biased linear regressors. Two major classes are identified: Bayesian estimators and restriction regressors. Both classes are useful for process applications, but the latter lack much of the unifying theory of the former. The properties of restriction regressors are examined and the characteristics of a "good" restriction regressor are expressed as a null hypothesis. Based on this characterization of "good," a novel restriction regressor is developed directly from the classical statistical concept of significance. This new method is called Significance Regression (SR). For scalar output problems, the popular Partial Least Squares algorithm (PLS) is a method for computing SR. For multiple output problems, SR yields a novel algorithm and PLS is sub-optimal.

SR is described using linear operator theory; this description allows facile generalization. SR is generalized for measurement error models and for robust regression methods. Analysis of these two extensions in turn provides insight an area that is currently dominated by useful heuristics with sparse mathematical grounding: scaling the data.

The theoretical results are illuminated by a variety of examples. First, several of the key points of the study are examined via simulation. Next, restriction regression is used in a robust inferential controller for a packed-bed react and for the modeling of cellular metabolism. Recommendations for implementing significance regression and suggestions for future research are provided.

Contents

Acknowledgments v
Abstract vii
List of Figures xii
List of Tables xiii
List of Algorithms xv
Nomenclature xvi
Preface xix

I Introduction **1**

1 Introduction **2**
1.1 The Problem 2
1.2 Two General Approaches for Treating Collinearity 5
1.3 Thesis Overview 7

II Development of Improved Linear Regression **9**

2 Biased Linear Regression **10**
2.1 Introduction 10
2.2 Measures of "Improvement" 12
2.3 Properties of "Restriction Regressors" 14
2.4 Issues for Constructing Restriction Regressors 18
2.5 Appendix: Reduced Variance of "Restriction Regressors" 20

3 Significance Regression for Scalar Output Problems **22**
3.1 Derivation 22
3.2 Choosing n_d 25
3.3 Some Properties of SR 30
3.4 Why not use the PRESS? 32
3.5 Conclusion 33
3.6 Appendix: Equivalence of SR and PLS for the Scalar Output Case . . 34

4 Linear Operator Description of Significance Regression **38**
4.1 Introduction 38
4.2 The General Null Hypothesis 39

4.3 Regression 40
4.4 Factor Analysis 41
4.5 A System Identification Illustration 44
4.6 Summary 46

III Generalizations and Improvements for Significance Regression 47

5 Multiple Output Significance Regression 48
5.1 Introduction 48
5.2 Regression 49
5.3 Factor Analysis 51
5.4 Relationship to Partial Least Squares 54
5.5 Tensor Problems 55
5.6 Summary 56

6 Measurement Error Model Significance Regression 57
6.1 Introduction 57
6.2 Significance Regression Method 59
6.3 Classical Regression Model Methods 62
6.4 Summary 65
6.5 Appendix: Proof of Claims 66

7 Robust Significance Regression 72
7.1 Introduction 72
7.2 Weighted Least Squares Significance Regression 73
7.3 Robust Regression 76
7.4 Robust Significance Regression 81
7.5 Summary 82

8 Scaling for Significance Regression 84
8.1 Introduction 84
8.2 Scaling Objectives 84
8.3 Aspects of Scaling 86
8.4 Removing the Effect of Units 87
8.5 Accounting for Heteroscedasticity 88
8.6 Mitigating Input Measurement Errors 90
8.7 Conclusions 92

IV Process Applications 93

9 Synthetic Examples 94
9.1 Introduction 94
9.2 Vector Outputs 96

9.3 The Measurement Error Model 99
9.4 Data with Outliers 102
9.5 Scaling 104
9.6 Conclusions 107

10 Composition Inference for Control of a Packed Bed Reactor 109
10.1 Introduction 109
10.2 Brosilow's Inferential Control 113
10.3 SR Prediction 117
10.3.1 SR Predictor 117
10.3.2 Uncertainty Description 121
10.3.3 Design of the Dynamic Compensator 124
10.4 Closed-Loop Experiments 125
10.4.1 Experiments for Brosilow's Scheme 126
10.4.2 Experiments Using the SR Inferential Scheme 129
10.5 Conclusions 131
10.6 Appendix: Nomenclature Unique to Chapter 10 134

11 Metabolic Control Analysis 135
11.1 Introduction 135
11.2 Metabolic Modeling 136
11.2.1 The General Metabolic Modeling Problem 137
11.2.2 The Specific Metabolic Modeling Problem 140
11.3 Numerical Results 142
11.4 Discussion of Key Issues 144
11.5 Conclusions 149

V Summary, Recommendations, and Future Work 150

12 Summary, Recommendations, and Future Work 151
12.1 Summary of Contributions 151
12.2 Recommendations 153
12.3 Suggestions for Future Work 154

Appendix 159

A PLS Review 158

B Routines To Generate Synthetic Data 162
B.1 Vector Output Examples 162
B.2 MEM Examples 162
B.3 Vector Output Examples with Outliers 163
B.4 Scaling Examples 163
B.5 The Random Example Generation Routines 164

B.5.1 Routine to Generate Random Regression Problems 164
B.5.2 Routine to Generate Random Regression Problems with Outliers 165
B.5.3 Routine to Generate "Exponential" Random Numbers 167

C Raw Data from Packed Bed Reactor Experiments **169**

D Metabolic Model for MCA Study **172**
D.1 Cellular Metabolism Model 172
D.2 Black Frame Analysis of the Cell Metabolism Model 175
D.3 Perturbations Used with the Cellular Metabolism Model 177

Bibliography **179**

List of Figures

1.1 Plot of the two temperatures for the four samples in Table 1.1 4

10.1 The time response of concentration to step in recycle flowrate for concentration. 120
10.2 Block structure for steady-state performance analysis for the SR inferential scheme 123
10.3 Response of the manipulated variable to step decrease in wall temperature. 128
10.4 Response of controlled variables to step decrease in wall temperature using Brosilow's inferential scheme. 129
10.5 Response of controlled variables to step decrease in wall temperature using SR inferential scheme. 130
10.6 Comparison between exit concentration predicted by SR predictor to concentration measured by GC. 131

11.1 Metabolic processes included in a cell mass model developed by Schlosser and Bailey[81]. 139

List of Tables

1.1 Temperature and composition data for a hypothetical chemical reactor. The data are expressed as deviations from a nominal operating point. 3

9.1 Comparison of PLS and SR using cross-validation over 1,000 examples of synthetic data. 97
9.2 Comparison of the 90 % significance test and cross-validation over 1,000 examples of synthetic data. 97
9.3 Comparison of the "over-simplified" significance test and cross-validation over 1,000 examples of synthetic data. 97
9.4 Comparison of MEM-based methods over 1,000 examples of synthetic data. 100
9.5 Comparison of asymptotically unbiased estimator versus least-squares estimator over 1,000 examples of synthetic data. 100
9.6 Comparison of SR-MEM and SR methods over 1,000 examples of synthetic data. 100
9.7 MEM prediction performance over 1,000 examples of synthetic data. 100
9.8 Comparison of OLS and M-est over 1,000 examples of synthetic data. 106
9.9 Comparison of SR-Robust and SR using significance testing over 1,000 examples of synthetic data with outliers. 106
9.10 Comparison of SR and SR-Robust using cross-validation over 1,000 examples of synthetic data with outliers. SR-Robust used "robust" PRESS. 106
9.11 Comparison of scaling regimes for treating input measurement errors when using an "over-simplified" test over 1,000 examples of synthetic data. 106
9.12 Comparison of scaling regimes for treating input measurement errors when using cross-validation over 1,000 examples of synthetic data. . . 108
9.13 Comparison of the *ad hoc* heteroscedastic methods over 1,000 examples of synthetic data. 108

10.1 Steady-state experiments using Brosilow's and SR predictors 116
10.2 Uncertainty bounds for gains, time constants and delays for the transfer functions relating manipulated variables to the predicted concentration using SR and the interpolated maximal bed temperature. 122
10.3 Closed-loop inferential control experiments 127

11.1 Experimental perturbations used in the MCA study. 143

11.2 "True" concentration control coefficients (*ccc*'s) 143
11.3 Concentration control coefficients (*ccc*'s) computed using OLS. . . . 143
11.4 Concentration control coefficients (*ccc*'s) computed using PLS. 143

C.1 Raw data for the 14 experiments used to build the PLS predictor for the methanation reactor. 170
C.2 Raw temperature data ($^{\circ}C$) for the 14 experiments used to build the PLS predictor for the methanation reactor 171

D.1 Experimental perturbations used in Schlosser and Bailey [81] 178
D.2 Nominal and deviation values "measured" from the nonlinear model for the perturbations in Table 11.1 178

List of Algorithms

3.1 Significance regression for scalar output problems 25
3.2 Helland's PLS algorithm for scalar output problems 34

4.1 General significance regression 41
4.2 General SR factor analysis 42

5.1 SR-Factor. Factor analysis for vector outputs 53

7.1 SR-WLS. Weighted least squares significance regression 74
7.2 M-estimation robust regression 79
7.3 SR-Robust. Robust significance regression 81

12.1 Recommended significance regression procedure 153

A.1 Generation of PLS loading vectors 160

Nomenclature

In general, script letters represent Hilbert spaces, capital letters represent matrices, lower case letters represent column vectors, and Greek letters represent scalars. Estimates are denoted by a tilde, "~". The dimensions of matrices are denoted by subscripted n's.

matrices

D_1	$n_i \times n_i$	is the symmetric non-singular "input" scaling matrix. See equation 8.1
D_2	$n_i \times n_i$	is the symmetric non-singular "output" scaling matrix. See equation 8.2
E	$n_s \times n_i$	is the measurement noise corrupting the input data. $E^T = [e_1 \;\; e_2 \;\; \ldots \;\; e_{n_s}]$ See equation 6.2.
I	as appropriate	is the identity matrix.
M	$n_s \times n_s$	is matrix of weights for the weighted least squares objective function. See equation 7.3
P	$n_s \times n_s$	is the output error covariance matrix.
R	$n_i \times n_o$	is the "true" regression matrix for vector output problems. See equation 5.1.
T	$n_s \times n_i$	is the "true" input data. See equation 6.2.
W	$n_i \times n_w$	is the matrix whose range defines the search space for $\tilde{b}$. See equation 2.8.
V	$n_i \times n_i$	is the estimate of the variance matrix used in algorithm 6.1. See equation 6.11.
V_{ideal}	$n_i \times n_i$	is the estimate of the variance matrix using the unknown quantity r. See equation 6.8.
X	$n_s \times n_i$	is the input data; each row corresponds to one input sample. Thus, $X^T = [x_1 \;\; x_2 \;\; \ldots \;\; x_{n_s}]$.
Y	$n_s \times n_o$	is the output data for vector output problems. Each row corresponds to one output sample. Thus, $Y^T = [y_1 \;\; y_2 \;\; \ldots \;\; y_{n_s}]$.
Σ	$n_i \times n_i$	is the input measurement error covariance matrix. See (A7), page 58.

vectors and scalars

$\tilde{b}$	$n_i \times 1$	is the biased estimate of r. See equation 2.8 .
e	$n_s \times 1$	is the measurement noise corrupting output. See equations 2.1.
e_j	$n_i \times 1$	is the measurement noise corrupting the jth input data sample in the measurement error model See equation 6.2.
p	$n_i \times 1$	is the least-squares optimal predictor. See equation 6.36.
q_1	$n_w \times 1$	is the component of r in the search space. $q_1 = W^T r$.
q_2	$n_i - n_w \times 1$	is the component of r orthogonal to the search space. $q_2 = W^{\perp T} r$.
r	$n_i \times 1$	is the "true" regression vector. See equations 2.1 and 6.2.
$\tilde{r}$	$n_i \times 1$	the Ordinary Least Squares (OLS) regressor. See equation 2.4.
$\tilde{r}_{\text{MEM}}$	$n_i \times 1$	is the asymptotically unbiased estimate of r for the measurement error model. See equation 6.3
v	varies$\times 1$	is a vector locally defined. Any given v may or may not relate to any other v.
x_j	$n_i \times 1$	is the jth input data sample.
y	$n_s \times 1$	is the measured output data for scalar output problems. See equations 2.1 and 6.1.
$\tilde{y}$	$n_s \times 1$	is the regression prediction for scalar output data.
y_i	$n_o \times 1$	is the ith output data sample for vector output problems.
z_i	$n_s \times 1$	is the vector produced by the ith input for all samples. Thus, $X = [z_1 \ \ldots \ z_{n_i}]$.
ψ_j	scalar	is the jth component of y. $y = [\psi_1, \ldots, \psi_{n_s}]$.
$\tau(w, y)$	scalar	is the test statistic for w and a given y. See equation 3.2
$(\tau_i^{opt}(y))^2$	scalar	is the maximum of $\tau^2(w, y)$ for $w \in \text{Range}(W_{i-1}^{\perp})$.
$\tau(S(\cdot), X, Y)$	scalar	is the test statistic for $S(\cdot)$ and a given X and Y. See equation 4.4

dimensional descriptors

n_d	is the number of "significant subspaces" to be generated.
n_i	is the number of inputs.
n_o	is the number of outputs.
n_s	is the number of samples.
n_p	is dimension of the allowable space in which to search for further w_i^{opt}. For scalar output problems, $n_p = n_i - i + 1$.
n_w	is the rank of W.

operators

$\|\cdot\|$	is the absolute value.
$\|\|\cdot\|\|$	is the Euclidean norm. $a = \sqrt{< a, a >}$.
$\|\|\cdot\|\|_2$	is the matrix 2-norm. $\|\|A\|\|_2 = \sqrt{\max_{v \in \Re^{n_i}} v^T A^T A v}$.
$\|\|\cdot\|\|_F$	is the Frobenius (matrix Euclidean) norm. $\|\|A\|\|_F = \sqrt{< A, A >} = \sqrt{\sum_{i,j} a_{i,j}^2}$, where $a_{i,j}$ are the components of A.
$[W \ \|V]$	is the matrix formed by placing W and V side-by-side.
$< \cdot, \cdot >$	is the inner product. For matrices A and B, $< A, B > = \mathrm{Tr}(AB^T)$.
$\mathcal{E}(\cdot)$	is the expectation.
MSE$(\cdot)$	is the Mean Square Error. See equation 2.9.
PRESS$(\cdot)$	is the PRedicted Error Sum of Squares. See equations 2.10 and 2.11.
Pr$\{event\}$	is the probability of *event* occurring.
Range$(\cdot)$	is the range of an operator. For a matrix, the range is the span of the column vectors.
Rank$(\cdot)$	is the dimension of the range of an operator.
Span$(\cdot)$	is the space defined by all linear combinations of the elements in a set.
Tr$(\cdot)$	is the trace, the sum of the diagonal elements of a matrix.
Var$(\cdot)$	is the variance.
$\otimes$	is the Cartesian (orthogonal) product when applied to spaces and is the outer product when applied to vectors. For example, $\Re^2 \otimes \Re^3 = \Re^{2\times 3}$.
$\oplus$	is direct (orthogonal) sum. For example, $\Re^2 \oplus \Re^3 = \Re^5$ if the two spaces are complementary.

spaces

$\mathcal{I}$	is the input space.
$\mathcal{M}$	is the space of linear operators mapping from $\mathcal{I}$ to $\mathcal{O}$. In this study, $\mathcal{M} = \mathcal{O} \otimes \mathcal{I}$
$\mathcal{O}$	is the output space.
$\mathcal{F}$	is the set of all allowable rank one linear operators. $\mathcal{F} \subseteq \mathcal{M}$.
$\mathcal{F}_i$	is the space spanned by the i-1 previously determined "significant subspaces."
$\mathcal{F}_i^\perp$	is the set of all possible rank one linear operators orthogonal to the i-1 previously determined "significant subspaces." $\mathcal{F}_i^\perp \oplus \mathcal{F}_i = \mathcal{F}$

Preface

While the author does desire to complete his studies, this was not the primary motivation for this thesis. The primary goal was to use the common language of statistics to understand a successful chemometric method and to communicate this method to other areas of applied statistics such as econometrics, technometrics, and biometrics. In doing so, this work builds a framework by which myriad restriction regressors can be described, related, compared, and evaluated. This work is also a user's guide for those who wish to employ the recently developed significance regression algorithm. Readers for whom this is the primary objective may wish to go directly to the recommendations of section 12.2.

An intended implication of this work is the removal of the "hard modeling" vs. "soft modeling" dichotomy from the debate surrounding the successful but heuristic partial least squares method. The spread of PLS from chemometrics has been slowed by "hard modelers," those who insist on rigorous statistical treatments, being uncomfortable with the sensible but *ad hoc* motivations supporting PLS. (To be fair, work such as Helland's and Höskuldsson's can hardly be termed *ad hoc*, but the debate continues all the same). On the other side of the fence, the "soft-modelers" have been convinced of PLS's effectiveness for almost half of a decade and have been impatient with their "hard" colleagues. The discussion following Stone and Brooks's recent article [88] is a good example of the debate. In this work, PLS is derived directly from the "hard modeling" concept of statistical significance. In this sense, the "soft modelers" were right all along: PLS really is a "good" method for which the theory is just now catching-up with practice. On the other hand the "hard modelers" were also correct. Translating PLS to new problem domains and improving PLS is much

clearer when using a rigorous framework. With this dialectic resolved, the author looks forward to new synergisms between "hard" and "soft" modelers.

Many important regression results have been omitted; this work is not a general survey of regression. Numerous successful (and rigorous) regressors are mentioned only in passing and are not analyzed. Also, no new probabilistic tools are developed for probing the abstract foundations of regression problems; this work uses ideas familiar to any with a solid grounding in statistics. And lastly, this thesis is not a compendium of the author's graduate experience; numerous successful ideas and results not germane to the purpose of this thesis have been omitted.

Part I

Introduction

Chapter 1

Introduction

1.1 The Problem

Fitting straight lines to data constitutes one of the most basic problems in science. One of the early experiences of all scientists is plotting X-Y data and using a ruler to "fit" the data. The problem has a long, proud history: Gauss published the least-squares method for fitting straight lines in 1809 [19] and applied the method to the triangulation of Hannover in 1826. For this surveying problem, "the least squares method served notoriously well [88]." Yet the least-squares method for linear models, termed ordinary least squares (despite it's extraordinary pedigree and scope of application), has numerous deficiencies. The following brief example illustrates one such problem.

When operating a chemical reactor, one is typically interested in the outlet concentration of the chemical species being produced; unfortunately, precise composition measurements are typically expensive and unreliable. Temperature measurements, which are known on physical grounds to be closely related to the outlet concentrations, are usually readily available, cheap, and reliable. One would like to be able to determine compositions from temperatures. To aid such an effort one could gather

Temperature No. 1	Temperature No. 2	Composition
1.1	1.0	2.26
−1.3	−1.2	−2.44
0.1	0.0	0.06
−2.3	−2.4	−4.52

Table 1.1. Temperature and composition data for a hypothetical chemical reactor. The data are expressed as deviations from a nominal operating point.

data on temperatures and concentrations from the reactor at various operating conditions; such data for a hypothetical reactor are shown in Table 1.1. Armed with data, one follows Gauss's lead by first assuming one's data conforms to the "classical" model

$$y = Xr + e \tag{1.1}$$

where

$$X = \begin{bmatrix} 1.1 & 1.0 \\ -1.3 & -1.2 \\ 0.1 & 0.0 \\ -2.3 & -2.4 \end{bmatrix} \quad \text{and} \quad y = \begin{bmatrix} 2.26 \\ -2.44 \\ 0.06 \\ -4.52 \end{bmatrix}. \tag{1.2}$$

The "noise" is described by e, while r represents the "linear model" or "linear predictor" one desires. Intuitively, one might expect that for this data $r = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ might be reasonable. Performing the regression via Gauss's "notoriously successful" OLS method produces $\tilde{r} = \begin{bmatrix} -11.0 \\ 12.2 \end{bmatrix}$. Not only does $\tilde{r}$ seem intuitively incorrect, $\tilde{r}$ deviates the model used to generate the data (any r whose elements sum to 2) by 40% ! For this problem, OLS is notorious, but not successful. What went wrong?

Inspecting Figure 1.1 reveals the problem. In this figure the two temperatures measurements for each sample are plotted; each point represents one row of X. One immediately sees that the temperature samples themselves all lie near a line; the input data are "collinear." The collinearity results from the fact that the input

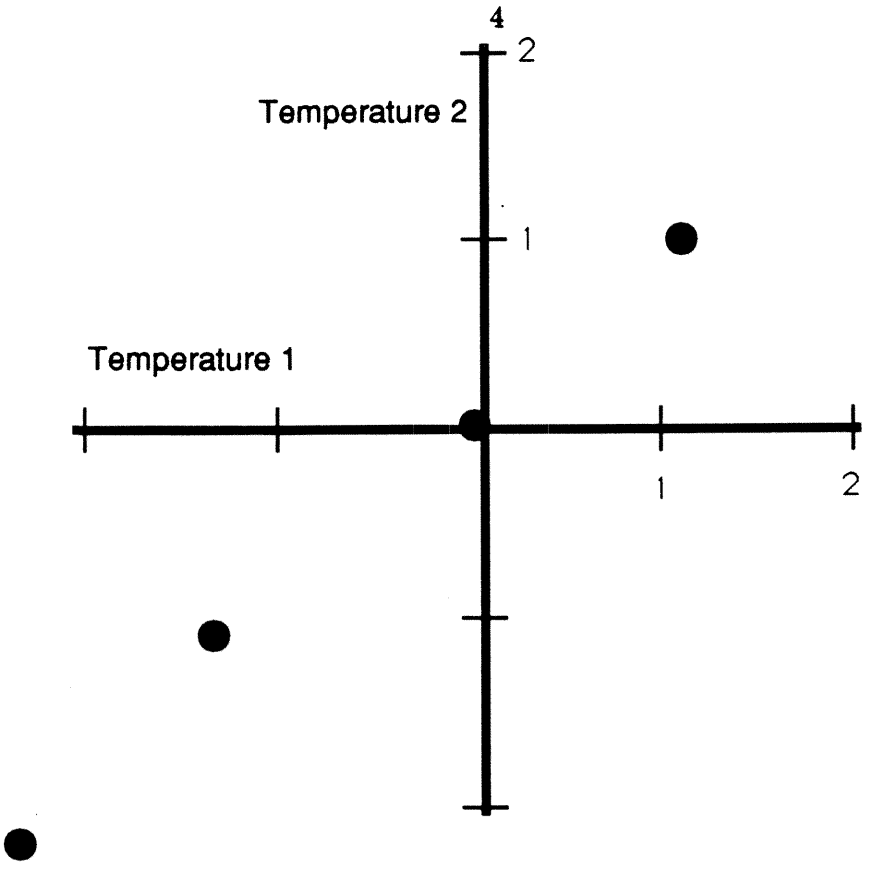

Figure 1.1. Plot of the two temperatures for the four samples in Table 1.1

(temperature) data are strongly correlated. The effect illustrated here is a general one: OLS will be unreliable when the input data are collinear (correlated). Nature and scientists alike rarely practice orthogonal experimental design, so many multivariable data sets exhibit input correlations and call for "improved" regression methods .

Before rushing to heed this call, two terms need to more clearly defined: regression and collinearity. Francis Galton, a noted scientist of the prior century, believed that the laws of inheritance caused population extremes to "regress toward the mean." To verify this, the English statistician Karl Pearson performed a linear fit between the heights of fathers and sons to test Galton's supposition; thereafter "regression" became associated with building linear models from data [76, p. 256]. While the study heredity is an important application of statistics, this work takes a more general

viewpoint and defines regression to be the process of either estimation or prediction. Estimation is the effort to identify the "true" but unknown parameters of the assumed model from the data; prediction is the use of data to form a predictor of outputs from a function of inputs. While estimation and prediction are closely related, one should keep the distinction clear in one's mind. The two problems usually call for different (albeit similar) methods.

The word "collinearity" is used in this thesis in the statistical sense. Strictly speaking, the temperatures in the above example are not collinear: the columns of X are linearly independent. However, as Figure 1.1 revealed, the temperatures are "nearly-collinear" — the data lie near a line. Throughout this treatise "collinear" will mean that the input data lie near hyper-plane whose dimension is less then n_i, the number of inputs. This "nearness" can be effectively quantified by the condition number of X; the condition number is defined as the square root of the ratio of the maximum and minimum eigenvalues of $X^T X$ [21]. Thus, problems with correlations among the inputs lead to "ill-conditioned" data (data with "large" condition number) and are considered collinear problems.

1.2 Two General Approaches for Treating Collinearity

To address the collinearity issue, consider again the ordinary least-squares (OLS) regressor

$$\tilde{r} = \arg\min_{v}(y - Xv)^T(y - Xv). \tag{1.3}$$

One could consider changing the objective function, as discussed in chapter 7, but most "improvements" to OLS result from recognizing that $\tilde{r}$ is too "free" to drive down the least-squares objective function in "unreasonable" ways, and thus some

type of condition needs to be added to equation 1.3 to constrain this "freedom." The two general methods for constraining an optimization problem are soft constraints and hard constraints. When one wishes to keep $\tilde{r}$ "near" some region or "away" from some other region of the search space, but the boundaries are not precise, one adds a penalty functional $f(\cdot)$ to equation 1.3 to encode this soft constraint:

$$\tilde{r} = \arg\min_{v}(y - Xv)^T(y - Xv) + f(v). \tag{1.4}$$

However, if the boundaries of the region to which $\tilde{r}$ should be restricted, say $\mathcal{R}$, are known precisely, then one has an optimization problem with hard constraints:

$$\tilde{r} = \arg\min_{v \in \mathcal{R}}(y - Xv)^T(y - Xv). \tag{1.5}$$

In this study regressors related to equation 1.4 are called "Bayesian regressors" because these regressors can be derived directly from a Bayesian viewpoint: the penalty function encodes one's prior beliefs about a problem. A large number of "improved" regressors, including ridge regression [34] belong to this class. These regressors now have much unifying theory [24] and can be considered reasonably well understood [59]. Regressors related to equation 1.5 include stepwise regression [14], principal components regression (PCR) [45], and partial least squares (PLS) [101]; these regressors have been as conspicuously successful as the Bayesian regressors. However these regressors lack much the unifying theory of their Bayesian counterparts. In some important cases, even the description of an unbiased regressor remains an open problem [90]. This thesis seeks to move the "hard constraint" regressors a little closer to "reasonably well understood" status. To do so, the work focuses on the most common form of this regressor:

$$\tilde{r} = \arg\min_{v \in \{\text{linear subspace}\}}(y - Xv)^T(y - Xv). \tag{1.6}$$

Such regressors will be called "restriction regressors" in this work because the regressor is restricted to be in linear subspace. Stepwise regression, PCR, and PLS are among the better known restriction regressors. While most recognize that these methods are related [20, 31] and some attempts have been made to unify these approaches via elaborate objective functions [58, 88], much remains to be discovered about the relations between the various restriction regressors and the relative strengths of the different methods.

1.3 Thesis Overview

This thesis first examines general properties of biased regressors. After a brief review of regression, the estimation and prediction properties of restriction regressors are analyzed. Based on this analysis a key property for a restriction regressor to be "good" is presented as a null hypothesis. A test statistic is derived to test this null hypothesis. Defining a method to maximize "good" in the guise of the test statistic leads to the derivation of a new restriction regressor, the significance regressor (SR). To ease generalization, the SR method is defined for both regression and factor analysis using linear operator theory.

With restriction regressors analyzed and SR defined, the work turns to addressing the needs of practitioners. First elucidated are vector output problems. As shown in this analysis PLS is similar to the SR factor analysis method but sub-optimal for regression. Next, one of the more glaring assumptions of Gauss's approach is interrogated: what are the consequences if the inputs (X) *do* have measurement errors? The SR method is applied to the measurement error model (MEM) to arrive at a new algorithm for these problems. Moreover, the performance of methods based on the classical model (equation 1.1) is examined for use with the MEM. After the issues of measurement errors are illuminated, the difficulties of robustly handling outliers are spotlighted. A robust SR algorithm is derived that maintains the robustness prop-

erties of M-estimators. In the course of this derivation a SR algorithm for weighted least squares is also derived. From here the inquiry reveals the skeleton in the closet of most restriction regressors: scaling. In examining the scaling problem one is reminded that restriction regressors differ from Bayesian regressors only in the manner in which the constraints are posed; determining the scaling relies heavily on the practitioner's *a priori* beliefs. However theoretical examination of SR does provide some guidance.

After inspecting the analysis of problems faced by practitioners this work next looks into real problems. First, the points and claims made in the course of analysis are illustrated via numeric simulation. Next, restriction regression is used for the robust inferential control of a packed bed reactor. The resulting robust inferential controller significantly outperforms a Kalman-filter based controller. Lastly, smaller but considerably more complicated chemical reactors are investigated: restriction regression is used to build a metabolic model to aid genetic engineering. Here the methods that were used previously by genetic engineers were simply inadequate; SR provides the tools to carry the analysis forward. More interesting than the success of SR metabolic modeling were the unmet challenges: this difficult problem highlighted both the strengths and shortcomings of the current SR framework.

Drawing from this practical experience, the theoretical results are compiled into a general procedure for performing restriction regression. The procedure incorporates the collinearity, measurement error, and robustness results produced in this study but is less complicated than some currently popular restriction regressors. Finally, glimpses are given of several promising avenues of future research.

Part II

Development of Improved Linear Regression

Chapter 2

Biased Linear Regression

2.1 Introduction

This chapter examines the "classical" linear regression model

$$y = Xr + e, \tag{2.1}$$

where $X \in \Re^{n_s \times n_i}$ and $y \in \Re^{n_s}$ are known, $r \in \Re^{n_i}$ is an unknown vector, and $e \in \Re^{n_s}$ is an unobservable error vector. Unless otherwise stated, $\mathcal{E}(e) = 0$ and $\mathcal{E}\left(ee^T\right) = \sigma_e^2 I$ where $\mathcal{E}(\cdot)$ denotes the expectation. Moreover, this work will assume that X and y have been "centered," so that the full model is actually

$$(y - v_0 \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix}) = (X - \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix} x_0^T)r + e. \tag{2.2}$$

The "centering" variables x_0 and v_0 represent either pre-specified nominal values or the sample means of X and y respectively. While Moran has shown that if X consists of samples from a multivariate distribution use of the sample means can

cause unreliable estimation [65], this work will proceed using "mean centering" when a relevant "nominal point" is not available. "Mean centering" is widely used, and moreover, frequently correct [86].

For the "classical" model (equation 2.1) the minimum-variance unbiased estimate of r has long been known to result from ordinary least-squares (OLS) regression, namely

$$\tilde{r} = \arg\min_{v \in \Re^{n_i}} (y - Xv)^T (y - Xv) \tag{2.3}$$

$$= (X^T X)^{-1} X^T y. \tag{2.4}$$

Equation 2.4 assumes $(X^T X)$ is non-singular; this assumption is maintained throughout this thesis. The variance of $\tilde{r}$, $\mathrm{Var}(\tilde{r}) = \sigma_\epsilon^2 (X^T X)^{-1}$, can be unacceptably large, especially if X is "ill-conditioned." If the condition number [21] is "large," then at least one direction in $(X^T X)^{-1}$ (relatively) magnifies σ_ϵ^2 and causes "large" variances in $\tilde{r}$. Thus collinear input data tend to produce unreliable regressors when OLS is used.

A natural approach for improving upon OLS is to "discourage" various values of $\tilde{r}$ by adding a quadratic penalty function to the optimization:

$$\tilde{b} = \arg\min_{v \in \Re^{n_i}} (y - Xv)^T (y - Xv) + v^T A v \tag{2.5}$$

$$= (X^T X + A)^{-1} X^T y \tag{2.6}$$

where $A \in \Re^{n_i \times n_i}$ is a positive semi-definite matrix. Such regressors will be called Bayesian regressors in this study. These regressors include the generalized ridge regressor [25], the mixed estimator [91], the linear minimax estimator [70], and the Bayes estimator for a zero mean *a priori* distribution for r [55]. See Gruber [24] for extensive analysis of these regressors and how they relate to one another.

A different but widely used approach is to impose linear equality constraints on

the solution to equation 2.3. This constrained optimization problem is

$$\tilde{b} = \arg \min_{v \in \text{Range}(W)} (y - Xv)^T (y - Xv) \quad (2.7)$$

$$= W(W^T X^T X W)^{-1} W^T X^T y \quad (2.8)$$

where $W \in \Re^{n_i \times n_w}$ and $W^T W = I$. Such regressors will be called restriction regressors in this study, are analyzed in general in section 2.3, and are the basis for this thesis. Clearly, the two methods are simply different sides of the same coin: if $A = \alpha W W^T$, then as $\alpha \to \infty$ equation 2.5 and equation 2.7 become equivalent. Although these two general approaches lead to very different regression methods, all methods that can be described by either optimization remain fundamentally linked.

Another important class of constraints are linear inequality constraints (*e.g.,* all parameters should be non-negative). This type of constraint arises frequently in modeling problems where the parameters have physical meanings. For instance, if one was performing regression on stress and strain data to compute a modulus of elasticity, one would typically only countenance positive numbers as "answers." Since these are by definition "hard" constraints, equation 2.7 is the natural departure point for such an investigation. Indeed, if linear inequality constraints are added to equation 2.7 then the minimization becomes the well-studied quadratic programming problem, for which efficient algorithms are widely available [15].

2.2 Measures of "Improvement"

In addition to being "optimal" in the sense of equation 2.3, OLS is an "unbiased" estimator: $\mathcal{E}(\tilde{r}) = r$. However, both Bayesian regressors and restriction regressors are "biased" estimators since in general $\mathcal{E}\left(\tilde{b}\right) \neq r$. Thus a more general quantitative measure of "improvement" than the variance is needed to compare the various

regressors. The two most common such measures are the mean squared error (MSE) and the predicted error sum of squares (PRESS). The MSE is defined as

$$\mathrm{MSE}(\tilde{r}) = \mathcal{E}\left((\tilde{r} - r)(\tilde{r} - r)^T\right) \tag{2.9}$$

while the PRESS is defined as

$$\mathrm{PRESS}(\tilde{b}) = (\tilde{y} - y)^T(\tilde{y} - y) \tag{2.10}$$

where $\tilde{y} = X\tilde{b}$. Typically the X and y used for evaluating the PRESS are different from the X and y used to create $\tilde{b}$. To increase understanding of the PRESS, consider computing the PRESS using a new set of data, X_{new}. Then

$$\mathcal{E}\left(\mathrm{PRESS}(\tilde{b})\right) = \mathcal{E}\left(\mathrm{Tr}(\ (\tilde{b} - r)^T X_{new}^T X_{new}(\tilde{b} - r)\ )\right) + \mathcal{E}\left(\mathrm{Tr}(e^T e)\right). \tag{2.11}$$

In this light, the PRESS and MSE are clearly related. As shown by Gruber [24] this relation can be made precise by means of Theobald's Theorem [92]. Let $\tilde{b}_1$ and $\tilde{b}_2$ be estimators of the parameter vector $r \in \Re^{n_i}$. Then

$$\mathcal{E}\left((\tilde{b}_1 - r)(\tilde{b}_1 - r)^T\right) \leq \mathcal{E}\left((\tilde{b}_2 - r)(\tilde{b}_2 - r)^T\right) \tag{2.12}$$

if and only if for every positive semi-definite matrix $A \in \Re^{n_i \times n_i}$

$$\mathcal{E}\left((\tilde{b}_1 - r)^T A(\tilde{b}_1 - r)\right) \leq \mathcal{E}\left((\tilde{b}_2 - r)^T A(\tilde{b}_2 - r)\right). \tag{2.13}$$

For any two matrices $A, B \in \Re^{n_i \times n_i}$, $A \geq B$ is $A - B$ is a positive semi-definite matrix. Therefore if

$$\mathrm{MSE}(\tilde{b}_1) - \mathrm{MSE}(\tilde{b}_2) \tag{2.14}$$

is a positive semi-definite matrix, then $\mathcal{E}\left(\text{PRESS}(\tilde{b}_1)\right) \geq \mathcal{E}\left(\text{PRESS}(\tilde{b}_2)\right)$; the converse may not be true. Additionally, one can see that using equation 2.11 involves assuming future inputs will be "similar" to past inputs. If $X_{new}^T X_{new}$ is descriptive of future X's, then equation 2.11 is clearly a measure of predictive performance. However, if the future inputs will have markedly different characteristics, or if the point of computing $\tilde{b}$ is to estimate r, then the PRESS may give misleading indications. Therefore the suitability of the PRESS or MSE will depend on the application. See, for example, chapters 12 and 13 of Ljung [56].

2.3 Properties of "Restriction Regressors"

A variety of widely-used regression methods are restriction regressors, including stepwise regression [14] and Principal Components Regression [84]. For stepwise regression each column would consist of unit vectors describing coordinate axes. For example if one chooses to use the second and third of three variables, then

$$W = \begin{bmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

For principal components regression W would be built from the principal components of X. These two methods are widely used and accepted. However, a so-called "soft-modeling" tool, the partial least squares (PLS) [101], is currently considered by many practitioners to provide superior prediction [20]. Though not immediately apparent in some descriptions, PLS is a restriction regressor [30]. Additionally, several restriction regressors have been put forward recently including continuum regression [88] and a related method due to Lorber and coworkers [58]. Many of these methods have been successful on specific examples. However, these methods have heuristic motivations and are difficult to describe, analyze, or compare using classical statistical methods. Additionally this lack of a statistical foundation has led many of these

methods to rely on cross-validation to answer questions traditionally analyzed by hypothesis testing. To lay the foundation evaluating these various methods this section outlines general properties shared by all restriction regressors; however, a comparative study of these various methods in light of these results is beyond the scope of the current effort.

Throughout the remainder of this section, assume that $W \in \Re^{n_i \times n_w}$ has been chosen independent of the error e and that the columns of W are orthonormal. Additionally, define $W^{\perp}$ such that $[W\ W^{\perp}]^T[W\ W^{\perp}] = I$ and $r = Wq_1 + W^{\perp}q_2$, where $q_1 \in \Re^{n_w}$ and $q_2 \in \Re^{n_i - n_w}$. One outstanding property of restriction regressors is

$$\mathrm{Var}(\tilde{r}) \geq \mathrm{Var}(WW^T\tilde{r}) \geq \mathrm{Var}(\tilde{b}). \tag{2.15}$$

One expects that since $\tilde{b}$ can only vary over a subspace of $\Re^{n_i}$ that the variance of a restriction regressor would be less than the variance of the OLS regressor. However, even when $\tilde{r}$ is projected into Range(W) the variance of the resulting regressor still dominates the variance of the corresponding restriction regressor. This key fact, which is the source of much of the MSE and PRESS advantage of restriction regressors, is proven in appendix 2.3. When the bias $(\mathcal{E}\left(\tilde{b} - r\right))$ is "small," the variance advantage of restriction regressors can lead to dramatic improvements over OLS.

Since $\tilde{b}$ is computed with the constraint $\tilde{b} \in$ Range(W), one expects $\tilde{b}$ to be a biased estimate of r. However, in addition to being unable to account for any of the $r \in$ Range($W^{\perp}$), $\tilde{b}$ has an additional bias due to trying to "stretch" in Range(W) to account for the "missing" r. This "stretching" occurs because the least squares estimator when the search space is restricted to a subspace of $\Re^{n_i}$ is not the projection of $\tilde{r}$ into the search space. Consider for the moment the case $\sigma_e^2 = 0$. Then

$$\begin{aligned} \tilde{b} &= \arg \min_{b \in \mathrm{Range}(W)} (y - Xb)^T(y - Xb) && (2.16) \\ &= W(W^TX^TXW)^{-1}W^TX^TXr, && (2.17) \end{aligned}$$

which is different from the projection of r into the search space, WW^Tr. In general $\mathcal{E}\left(\tilde{b}\right) = W(W^TX^TXW)^{-1}W^TX^TXr$. Thus, there will typically be two bias terms.

$$\begin{aligned}
\mathcal{E}\left(\tilde{b}-r\right) &= (W(W^TX^TXW)^{-1}W^TX^TX - I)r && (2.18)\\
&= (W(W^TX^TXW)^{-1}W^TX^TX - I)(Wq_1 + W^{\perp}q_2) && (2.19)\\
&= W(W^TX^TXW)^{-1}W^TX^TXW^{\perp}q_2 - W^{\perp}q_2. && (2.20)
\end{aligned}$$

The first bias term is due to the "stretching " of $\tilde{b}$ to account for variations in the output attributable to q_2 that can be partially described from Range(W), while the second bias term is the direct contribution of q_2. The two terms can never "cancel out" since they have complementary range spaces. However for the special case of PCR, $W^TX^TXW^{\perp} = 0$.

Consider the situation where $\tilde{b}$ has been constructed and a new input vector x_{new} is available. Computing the PRESS using x_{new} leads to

$$\begin{aligned}
\mathcal{E}\left((r^Tx_{new} + e_{new} - \tilde{b}^Tx_{new})^2\right) &= \left(x_{new}^T(W(W^TX^TXW)^{-1}W^TX^TX - I)r\right)^2\\
&\quad + x_{new}^TW(W^TX^TXW)^{-1}W^Tx_{new}\sigma_e^2 + \sigma_e^2 \quad (2.21)\\
&= \left(x_{new}^T(W(W^TX^TXW)^{-1}W^TX^TXW^{\perp}q_2 - W^{\perp}q_2)\right)^2\\
&\quad + x_{new}^TW(W^TX^TXW)^{-1}W^Tx_{new}\sigma_e^2 + \sigma_e^2. \quad (2.22)
\end{aligned}$$

Clearly, when $q_2 = 0$ the bias terms vanish and $\tilde{b}$ is superior to $\tilde{r}$. For prediction, even if there is a large bias term, the prediction bias will still tend to be small if x_{new} is not collinear with the bias. Thus restriction regressors are attractive for prediction problems even when one cannot assure that the bias is "small." One should note that requiring $x_{new} \in$ Range(W) does not assure unbiased prediction; there may still be bias arising from the "stretching" term.

Next, consider $\tilde{b}$ as a point estimator for r. In particular, evaluate

$$\begin{aligned}
\mathrm{Tr}\left(\mathrm{MSE}(\tilde{b})\right) &= \mathrm{Tr}\left(\mathcal{E}\left((r-\tilde{b})(r-\tilde{b})^T\right)\right) && (2.23)\\
&= \|(W(W^TX^TXW)^{-1}W^TX^TX - I)r\|^2 && (2.24)\\
&\quad + \sigma_e^2\mathrm{Tr}\left(W(W^TX^TXW)^{-1}W^T\right)\\
&= \|(W(W^TX^TXW)^{-1}W^TX^TXW^{\perp}q_2 - W^{\perp}q_2\|^2 && (2.25)\\
&\quad + \sigma_e^2\mathrm{Tr}\left(W(W^TX^TXW)^{-1}W^T\right).
\end{aligned}$$

As in the point prediction problem, two bias terms associated with q_2 appear. For the restriction estimator to have an MSE advantage over OLS, q_2 must be "small." Thus a successful restriction regressor (in the MSE sense) must strive to satisfy $r \in$ Range(W).

Move next to the problem of estimating an interval within which r^Tx_{new} lies with a certain probability. If one assumes that e is normally distributed, then one can use $\tilde{r}$ in the classical manner and declare that $\tilde{r}^Tx_{new} - r^Tx_{new}$ has a zero-mean normal distribution with variance $x_{new}^T(X^TX)^{-1}x_{new}\sigma_e^2$. Using, $\tilde{b}$, one can easily see that $r^Tx_{new} - \tilde{b}^Tx_{new}$ has a normal distribution with mean $x_{new}^TW(W^TX^TXW)^{-1}W^TX^TX - I)r$ and variance $x_{new}^TW(W^TX^TXW)^{-1}W^Tx_{new}\sigma_e^2$. Since r enters explicitly into the mean of the distribution, this cannot be used directly for a practical interval predictor. However if one is willing to conjecture that the bias is "small," then one can assume $r^Tx_{new} - \tilde{b}^Tx_{new}$ has a zero-mean normal distribution with variance $x_{new}^TW(W^TX^TXW)^{-1}W^Tx_{new}\sigma_e^2$. Under this assumption (which is equivalent to $q_2 = 0$), one has a smaller prediction interval using $\tilde{b}$ than when using $\tilde{r}$ since $(X^TX)^{-1} \geq W(W^TX^TXW)^{-1}W^T$, as shown in the appendix 2.3.

Lastly, turn to the problem of constructing a region of $\Re^{n_i}$ within which r lies with a certain probability. The classical approach is based on the observation that if

e is normally distributed then

$$\frac{(\tilde{r}-r)^T X^T X(\tilde{r}-r)}{\sigma_e^2} \tag{2.26}$$

has a χ^2 distribution with n_i degrees of freedom. One can adapt this approach to restriction regressors in a variety of ways, but the resulting confidence regions are generally more conservative. This outcome is not surprising. The traditional confidence bounds provide a region which contains r with a certain probability. A restriction regressor may choose a "better" point estimator from this region than the OLS point estimate, but the restriction regression process typically does not alter the original confidence region.

2.4 Issues for Constructing Restriction Regressors

Returning to the issue of building the biased regressor, the difficulty with equation 2.8 is that one must specify W. One approach for specifying W is to choose W so as to optimize the MSE:

$$W^{\text{MSE}} = \arg \min_{W \in \Re^{n_i}} \text{MSE}(\tilde{b}). \tag{2.27}$$

The solution is $W^{\text{MSE}} = \frac{r}{\|r\|}$ [2]. This "optimal" solution is interesting but not very useful because r is unknown, leaving the question of how to choose W unresolved. Still this "optimal" solution, as well as the properties discussed above, highlight the key feature one desires from a restriction regressor: that $q_2 = W^{\perp T} r = 0$.

Instead of computing the MSE-optimal W directly, one could try to build W incrementally by some other criterion in the hope of constructing a W such that $r \in \text{Range}(W)$. Consider for the moment an additional vector w such that $W^T w = 0$ and $\|w\| = 1$. One then asks: does using the search space Range($[W|w]$) lead to a $\tilde{b}$

with an MSE less than or equal to the MSE of $\tilde{b}$ using the search space Range(W)? The answer is "yes" if

$$(w^T r)^2 \geq \mathrm{Tr}\left([W|w] \left([W|w]^T X^T X)[W|w] \right)^{-1} [W|w]^T - W(W^T X^T X W)^{-1} W^T \right) \sigma_e^2. \tag{2.28}$$

The left hand side clearly reflects the bias removed from $\tilde{b}$ and the right-hand side reflects the variance added to $\tilde{b}$ by including Range(w) in the search space. Although equation 2.28 is not directly useful because r appears explicitly, it points the way: directions for which $w^T r = 0$ should not be included in the search space, and directions for which $w^T r$ is "large" and the variance is "small" should be included in the search space. These two observations provide the motivation for the particular restriction regressor developed in the next chapter.

2.5 Appendix: Reduced Variance of "Restriction Regressors"

Section 2.3 claims $\text{Var}(WW^T\tilde{r}) \geq \text{Var}(\tilde{b})$. The proof below is based directly on lemma 3.1 of Wahlberg[94].

Theorem 2.1 *For the model in equation 2.1 and any given W such that $W^TW = I$, $\text{Var}(WW^T\tilde{r}) - \text{Var}(\tilde{b})$ is always a positive semi-definite symmetric matrix.*

Proof. First, recall

$$\text{Var}(WW^T\tilde{r}) = WW^T(X^TX)^{-1}WW^T\sigma_e^2 \text{ and} \tag{2.29}$$

$$\text{Var}(\tilde{b}) = W(W^TX^TXW)^{-1}W^T\sigma_e^2. \tag{2.30}$$

Let Range(V) equal the null space of W^TX^TX. Then $W^TX^TXV = 0$ and $[W \mid V]$ is full rank. Next,

$$(X^TX)^{-1} = [W \mid V]\left([W \mid V]^TX^TX[W \mid V]\right)^{-1}[W \mid V]^T \tag{2.31}$$

$$= [W \mid V]\begin{bmatrix} (W^TX^TXW)^{-1} & 0 \\ 0 & (V^TX^TXV)^{-1} \end{bmatrix}[W \mid V]^T \tag{2.32}$$

$$= W(W^TX^TXW)^{-1}W^T + V(V^TX^TXV)^{-1}V^T. \tag{2.33}$$

Pre- and post-multiplying equation 2.33 by $WW^T\sigma_e$ yields

$$WW^T(X^TX)^{-1}WW^T\sigma_e^2 = W(W^TX^TXW)^{-1}W^T\sigma_e^2 + WW^TV(V^TX^TXV)^{-1}V^TWW^T\sigma_e^2 \tag{2.34}$$

which becomes

$$\text{Var}(WW^T\tilde{r}) - \text{Var}(\tilde{b}) = WW^TV(V^TX^TXV)^{-1}V^TWW^T\sigma_e^2. \tag{2.35}$$

Noting that $WW^TV(V^TX^TXV)^{-1}V^TWW^T$ is symmetric and positive semi-definite completes the proof. □

Since $\mathrm{Var}(WW^T\tilde{r})$ dominates $\mathrm{Var}(\tilde{b})$ by a positive semi-definite matrix, one may invoke other matrix results [38, page 471] and further state:

$$\mathrm{Tr}\left(\mathrm{Var}(WW^T\tilde{r})\right) \geq \mathrm{Tr}\left(\mathrm{Var}(\tilde{b})\right) \tag{2.36}$$

$$\|\mathrm{Var}(WW^T\tilde{r})\|_2 \geq \|\mathrm{Var}(\tilde{b})\|_2, \text{ and} \tag{2.37}$$

$$\|\mathrm{Var}(WW^T\tilde{r})\|_F^2 \geq \|\mathrm{Var}(\tilde{b})\|_F^2. \tag{2.38}$$

Chapter 3

Significance Regression for Scalar Output Problems

3.1 Derivation

The previous chapter discussed the estimation and prediction properties of restriction regressors. This chapter derives a specific regressor that is "optimal" in well-defined sense. The statistical framework underlying the derivation allows the development of significance tests as an alternative to cross-validation for determining $n_d = \text{Rank}(W)$. Moreover, the resulting regressor has additional useful properties beyond the properties in the previous chapter.

The linchpin of all these results is equation 2.28, which states clearly that if a particular subspace, described by w, is orthogonal to r then this subspace should be excluded from the W of the restriction regressor (equation 2.8). The idea translates directly into the null hypothesis

$$\mathcal{H}_0^1 : \qquad < r, w > = 0 \tag{3.1}$$

where $< r, w >$ is the vector inner product, $r^T w$. For this null hypothesis, a natural test statistic is

$$\tau(w, y) = \frac{< \tilde{r}, w >}{\sqrt{\text{Var} < \tilde{r}, w >}} \tag{3.2}$$

where

$$\begin{aligned} \text{Var} < \tilde{r}, w > &= \mathcal{E}\left((< \tilde{r}, w > - \mathcal{E}(< \tilde{r}, w >))^2 \right) && (3.3) \\ &= \mathcal{E}\left(< (X^T X)^{-1} X^T e, w > {}^2 \right) && (3.4) \\ &= w^T (X^T X)^{-1} w \sigma_e^2. && (3.5) \end{aligned}$$

Under the additional assumption that the errors are normally distributed, one can easily see that $\tilde{r}$ and $\tau(w, y)$ are normally distributed. If the noise variance σ_e^2 is unknown, then the unbiased estimate $\tilde{\sigma}_e^2 = \frac{1}{n_s - n_i}(y - X\tilde{r})^T (y - X\tilde{r})$ may be used instead; for normal errors, $\tilde{\sigma}_e^2$ arises from a χ^2 distribution. Since $\tau(w, y)$ varies as the error vector e projected into Range(X), while $\tilde{\sigma}_e^2$ varies as the error vector e projected into the orthogonal complement of Range(X), the two terms are independent. Throughout this chapter $\tilde{\sigma}_e^2$ can be used in place of σ_e^2 and the relevant distributions modified in the obvious manner. If $\tilde{\sigma}_e^2$ is used in equation 3.2 then $\tau(w, y)$ is associated with the Student's t-distribution. Using these distributions one can use classical significance testing procedures to evaluate $\mathcal{H}_0^1$ and to identify w that should be excluded from W.

As shown in section 2.3, one would like to identify w for which "$r^T w$ is 'large' and the variance is 'small.'" One can do this directly by computing w for which the absolute value of the $\tau(w, y)$ to be "large." In this sense the "most significant subspace" is described by

$$w_1^{opt}(y) = \arg \max_{w \in \Re^{n_i}} |\tau(w, y)|. \tag{3.6}$$

Since

$$\arg \max_{w \in \Re^{n_i}} |\tau(w,y)| = \arg \max_{w \in \Re^{n_i}} (\tau(w,y))^2, \tag{3.7}$$

$(\tau(w,y))^2$ will be used for the derivation. Equation 3.2 reveals that w_1^{opt} will not be unique; multiplying any w by a scalar will not affect the value of $(\tau(w,y))^2$. Still, the necessary condition for an unconstrained extremum

$$\nabla_w \ (\tau(w,y))^2 \Big|_{w_1^{opt}(y)} = 0 \tag{3.8}$$

must be met. Computing the gradient of $(\tau(w,y))^2$ gives

$$\nabla_w (\tau(w,y))^2 = \nabla_w \frac{(w^T \tilde{r})^2}{w^T (X^T X)^{-1} w \tilde{\sigma}_\epsilon^2} \tag{3.9}$$

$$= \frac{2\tilde{r}(w^T\tilde{r}) w^T (X^T X)^{-1} w \tilde{\sigma}_\epsilon^2 - 2(w^T\tilde{r})^2 (X^T X)^{-1} w \tilde{\sigma}_\epsilon^2}{(w^T (X^T X)^{-1} w \tilde{\sigma}_\epsilon^2)^2} \tag{3.10}$$

and applying equation 3.8 gives

$$w_1^{opt}(y) = X^T X \tilde{r} = X^T y. \tag{3.11}$$

As revealed in section 2.3, an important goal for any restriction regressor is to determine a W such that $r \in \text{Range}(W)$. Therefore W may need to have more than one column. Consistent with section 2.3, one searches for w that maximize $(\tau(w,y))^2$ and are orthogonal to the current $W_{i-1} = [w_1^{opt} | \ldots | w_{i-1}^{opt}]$. Then the ith "significant subspace" is described by

$$w_i^{opt} = \arg \max_{w \in \text{Range}(I - W_{i-1} W_{i-1}^T)} (\tau(w,y))^2. \tag{3.12}$$

Invoking the necessary condition for a constrained extremum yields

$$(I - W_{i-1} W_{i-1}^T) \quad \nabla_w \ (\tau(w,y))^2 \Big|_{w_i^{opt}(y)} = 0. \tag{3.13}$$

As shown in appendix 3.5, the PLS loading vectors satisfy equation 3.13. Thus one can use the PLS "soft modeling" algorithm to find n_d "significant vectors;" the issue of computing the solution to equation 3.13 has already been solved. While several different (but equivalent) PLS algorithms exist, Helland's algorithm [29] is the most straightforward and is used here.

Algorithm 3.1 (SR. Significance regression for scalar output problems)

$$\tilde{r} = (X^T X)^{-1} X^T y \tag{3.14}$$

$$W_0 = [0 \cdots 0]^T, \quad W_0 \in \Re^{n_i} \tag{3.15}$$

DO $i = 1, n_d$

$$v = (I - W_{i-1} W_{i-1}^T)(X^T X)^i \tilde{r} \tag{3.16}$$

$$= (I - W_{i-1} W_{i-1}^T)(X^T X)^{i-1} X^T y \tag{3.17}$$

$$w_i^{opt}(y) = \frac{v}{||v||} \tag{3.18}$$

$$W_i = [w_1^{opt} | w_2^{opt} | \cdots | w_i^{opt}] \tag{3.19}$$

END DO.

3.2 Choosing n_d

The above developments have assumed that n_d is known; however, in practice n_d needs to be determined. In the PLS context, the most popular method is cross-validation [98]. For any given n_d, $\tilde{b}$ is computed with a subset of the available data, and the PRESS is computed for that $\tilde{b}$ using the remainder of the data. Next the PRESS for different n_d's is compared to determine the "best" value of n_d. $\tilde{b}$ is then recomputed using all available data and the "best" n_d.

As discussed in section 2.2, the PRESS is an intuitively appealing measure when one is building predictors, but may not necessarily be the best measure for evaluating

estimates of r. The PRESS also has other potential drawbacks. If the data are sparse, withholding a portion of the data may be problematic. The technique of "one-out cross-validation" [87] largely overcomes sparseness problems by withholding a single sample in the data, computing a $\tilde{b}$, computing the PRESS for the withheld sample, and repeating this process for each sample in the data set. While this can be an effective approach for evaluating various values for n_d, the computational demands of one-out cross-validation can be excessive [57]. Clearly practitioners would benefit from the development of additional techniques for choosing n_d.

As shown in section 2.3, a useful condition for any restriction regressor to satisfy is $q_2 = 0$. This leads directly to the null hypothesis

$$\mathcal{H}_0^{2,i}: \qquad < r, w > = 0 \quad \forall w \in \text{Range}(I - W_{i-1}W_{i-1}^T) \tag{3.20}$$

for evaluating if $n_d = i - 1$. Let $\left(\tau_i^{opt}(y)\right)^2 = \left(\tau(w_i^{opt}, y)\right)^2$. If

$$\Pr\left\{ \; t^2 \leq \left(\tau_i^{opt}(y)\right)^2 \; \right\} \geq \alpha_{thresh}, \tag{3.21}$$

where α_{thresh} is some pre-specified significance threshold (90%, say) and t^2 is drawn from the distribution for $\left(\tau_i^{opt}(y)\right)^2$ when $\mathcal{H}_0^{2,i}$ holds, then $\mathcal{H}_0^{2,i}$ can be rejected and $n_d \geq i$.

Since W_{i-1} depends on y, evaluation of this distribution will often be involved. However, an approximate but useful distribution for $(\tau_i^{opt}(y))^2$ will now be developed for testing $\mathcal{H}_0^{2,i}$. The development rests on the key simplifying assumption that W_{i-1} was specified independent of e. Additionally, the remainder of this section assumes that the elements of e are independently identically normally distributed; if w is specified independently of y then $\tau(w, y)$ has a Gaussian distribution. Moreover, $\tau(w, y)$ has a unit normal distribution under $\mathcal{H}_0^1$.

We know from algorithm 3.1 that $w_1^{opt}(y) = X^T y \; / \; \|X^T y\|$. Turning attention to

the $(\tau_1^{opt}(y))^2$ that results from maximizing the significance criterion,

$$
\begin{aligned}
(\tau_1^{opt}(y))^2 &= \left(\tau(w_1^{opt}(y), y)\right)^2 && (3.22)\\
&= \frac{(y^T X(X^T X)^{-1} X^T y)^2}{y^T X(X^T X)^{-1} X^T y \sigma_e^2} && (3.23)\\
&= \frac{y^T X(X^T X)^{-1} X^T y}{\sigma_e^2} && (3.24)\\
&= \frac{\left((X^T X)^{\frac{1}{2}} r + (X^T X)^{-\frac{1}{2}} X^T e\right)^T \left((X^T X)^{\frac{1}{2}} r + (X^T X)^{-\frac{1}{2}} X^T e\right)}{\sigma_e^2}. && (3.25)
\end{aligned}
$$

Thus $(\tau_1^{opt}(y))^2$ has a non-central χ^2 distribution with n_i degrees of freedom and non-centrality parameter $r^T X^T X r$. Once again, if $\tilde{\sigma}_e^2$ is used in place of σ_e^2, then $(\tau_1^{opt}(y))^2$ has a non-central F distribution with $(n_i, n_s - n_i)$ degrees of freedom. Notice that when $r^T X^T X r$ is large, $(\tau_1^{opt}(y))^2$ has a small variance relative to $r^T X^T X r$. This small variance is reflected in $w_1^{opt}(y) = X^T X r + X^T e$. As $r^T X^T X r$ becomes large, $w_1^{opt}(y)$ is less influenced by e. When $\mathcal{H}_0^{2,1}$ holds ($r = 0$), then

$$
(\tau_1^{opt}(y))^2 = \frac{e^T X(X^T X)^{-1} X^T e}{\sigma_e^2}, \tag{3.26}
$$

and $(\tau_1^{opt}(y))^2$ has a χ^2 distribution with n_i degrees of freedom. Thus, under the null hypothesis $\mathcal{H}_0^{2,1}$, $(\tau_1^{opt}(y))^2$ has a familiar distribution; however, if a direction exists which strongly refutes $\mathcal{H}_0^2$, then SR will tend to identify this direction, and the subspace represented by $w_1^{opt}(y)$ will tend to be weakly affected by e.

Now consider the $(\tau_i^{opt}(y))^2$ resulting from determining $w_i^{opt} \in \text{Range}(I - W_{i-1} W_{i-1}^T)$. Assume for the moment that W_{i-1} is independent of e and let $\text{Range}(W_{i-1}^{\perp}) = \text{Range}(I - W_{i-1} W_{i-1}^T)$ and ${W_{i-1}^{\perp}}^T W_{i-1}^{\perp} = I$. Then

$$
w_i^{opt}(y) = W_{i-1}^{\perp} \left({W_{i-1}^{\perp}}^T (X^T X)^{-1} W_{i-1}^{\perp}\right)^{-1} {W_{i-1}^{\perp}}^T \tilde{r} \tag{3.27}
$$

and $(\tau_i^{opt}(y))^2$ has a non-central χ^2 distribution (as above) with $n_p = n_i - i + 1$ degrees of freedom and non-centrality parameter $r^T W_{i-1}^{\perp} \left({W_{i-1}^{\perp}}^T (X^T X)^{-1} W_{i-1}^{\perp} \right)^{-1} {W_{i-1}^{\perp}}^T r$. Once again, if $\mathcal{H}_0^{2,i}$ holds, then $(\tau_i^{opt}(y))^2$ has a χ^2 distribution, but if a direction strongly violates the null hypothesis, $w_i^{opt}(y)$ will be relatively unaffected by e.

These observations have tangible implications. If one wishes to evaluate $\mathcal{H}_0^{2,i}$ using $(\tau_i^{opt}(y))^2$, one can consider approximating the distribution of $(\tau_i^{opt}(y))^2$ with a χ^2 distribution with n_p degrees of freedom. In fact, such an approximation is valid in several asymptotic limits. Clearly, as the noise vanishes, the dependence of $w_i^{opt}(y)$ on e vanishes. That is

$$\lim_{\sigma_e^2 \to 0} \text{Span} \left(W_i^{opt}(y) \right) = \text{Range} \left([X^T X r | \ldots | (X^T X)^i r] \right). \tag{3.28}$$

Thus, when the noise is small enough, the independence assumption is justified. The independence assumption can also be justified when n_s is large. Consider again $w_1^{opt}(y)$ and the condition that the input data is persistently exciting ($\lim_{n_s \to \infty} \frac{1}{n_s} X^T X = A$ for some non-singular A). Then

$$\lim_{n_s \to \infty} w_1^{opt}(y) = \lim_{n_s \to \infty} \frac{\frac{1}{n_s} X^T y}{\left\| \frac{1}{n_s} X^T y \right\|} \tag{3.29}$$

$$= \lim_{n_s \to \infty} \frac{\frac{1}{n_s} X^T X r + \frac{1}{n_s} X^T e}{\left\| \frac{1}{n_s} X^T X r + \frac{1}{n_s} X^T e \right\|} \tag{3.30}$$

$$= \frac{Ar}{\|Ar\|}. \tag{3.31}$$

In this limit, $w_1^{opt}(y)$ is independent of e, and $(\tau_2^{opt}(y))^2$ has a non-central χ^2 distribution with $n_i - 1$ degrees of freedom. One can also show that all $w_i^{opt}(y)$ obey similar limits. Thus $(\tau_i^{opt}(y))^2$ has a non-central χ^2 distribution with n_p degrees of freedom in the limit of large n_s.

The above arguments have motivated using the independence assumption for computing the distribution of $(\tau_i^{opt}(y))^2$. However, before one risks using such an assump-

tion, one needs to know when and how the assumption will break down. When $r^T W_{i-1}^{\perp} \left({W_{i-1}^{\perp}}^T (X^T X)^{-1} W_{i-1}^{\perp} \right)^{-1} {W_{i-1}^{\perp}}^T r$ becomes comparable to n_p (the dimension of the search space, $n_p = n_i - i + 1$), the independence assumption will be incorrect for $w_j^{opt}(y)$ and $(\tau_j^{opt}(y))^2$ when $j > i$. Consider briefly the extreme case $r = 0$, $X^T X = I$. Then $(\tau_1^{opt}(y))^2 \sim \chi^2_{n_i}$. However the independence assumption completely breaks down for $w_2^{opt}(y)$: $(\tau_2^{opt}(y))^2 = 0$. This simplified example illustrates a larger point: as the independence assumption begins to break down, the earlier directions "steal" variance from later directions, and the correct distributions of $(\tau_i^{opt}(y))^2$ for later directions will have smaller tails than the distributions computed using the independence assumption.

Thus, one may usefully *approximate* the distribution of $(\tau_i^{opt}(y))^2$ with a χ^2 distribution with n_p degrees of freedom so long as the non-centrality parameter dominates the variance for earlier directions; encouragingly, these are precisely the directions which SR seeks. Moreover, the independence assumption breaks down in a known manner: the independence assumption leads to distributions with tails that are too heavy. When using the independence assumption with equation 3.21 to choose n_d, a test using the independence assumption will choose an n_d less than or equal to the n_d determined using that same test with the distribution that properly accounts for the dependence of $w_i^{opt}(y)$ on e.

Another implication of the observations stated above concerns the dependence of $W = [w_1^{opt} | \ldots | w_{n_d}^{opt}]$ on e. As discussed, SR violates the independence assumption essential for the derivation of the results of section 2.3. However, the w_i^{opt} for which the independence assumption that the distribution of $(\tau_i^{opt}(y))^2$ is independent of e is reasonable are the same w_i^{opt} for which the assumption that w_i^{opt} is independent of e is reasonable. Thus section 2.3 is approximately valid for SR for W consisting of w_i^{opt} where the non-centrality parameter dominates the variance.

For ease of computation, one may empirically "over-simplify" the above signifi-

cance test. When $\mathcal{H}_0^{2,i}$ and the independence assumption hold, n_s and n_p are "large," and a 50% significance threshold is used, the χ^2 test (or F-test) for $\mathcal{H}_0^{2,i}$ becomes:

$$\text{reject} \quad \mathcal{H}_0^{2,i} \quad \text{if} \quad (\tau_i^{opt}(y))^2 > n_p. \tag{3.32}$$

Even when the assumptions are not met, the ease of the "over-simplified" significance test may militate in favor of its use.

The significance tests developed here should not be viewed as a replacement for cross-validation, but as a complement. Often the two approaches will give similar determinations of n_d. However, the cross-validation techniques and significance tests rest on different assumptions and have varying computational needs. Significance tests will tend to impose less computational burden, but cross-validation will tend to be less impacted if the data deviate from the noise assumptions. Moreover, other approaches for determining n_d can be developed from the viewpoint developed in this chapter.

3.3 Some Properties of SR

Because of the statistical basis for SR, one can investigate properties for this regression method beyond those discussed in section 2.3. This section discusses further results specific to SR. Computing the expectation value of $W_i^{opt}(y)$ is involved. However, for the sake of determining the expected value of the search space, one can use the results in section 3.5 and state

$$\begin{aligned} \mathcal{E}\left(\text{Range}(W_i)\right) &= \text{Span}\left(\mathcal{E}\left([w_1^{opt}|\ldots|w_i^{opt}]\right)\right) && (3.33)\\ &= \text{Span}\left(\mathcal{E}\left([X^TX\tilde{r}|\ldots|(X^TX)^i\tilde{r}]\right)\right) && (3.34)\\ &= \text{Span}\left([X^TXr|\ldots|(X^TX)^ir]\right). && (3.35) \end{aligned}$$

Thus, $W_i^{opt}(y)$ provides an unbiased estimate of the "true" search space. Alternatively, consider the behavior of $W_i^{opt}(y)$ as n_s is increased. Assume that the input data is persistently exciting, that is $\lim_{n_s\to\infty} \frac{1}{n_s} X^T X = A$ for some non-singular A. For any w,

$$\lim_{n_s\to\infty} (\tau(w,y))^2 = \lim_{n_s\to\infty} \frac{(w^T\tilde{r})^2}{w^T(X^TX)^{-1}w\sigma_e^2} \tag{3.36}$$

$$= \lim_{n_s\to\infty} \frac{n_s(w^T\tilde{r})^2}{w^T(\frac{1}{n_s}X^TX)^{-1}w\sigma_e^2} \tag{3.37}$$

$$= \begin{cases} \infty & \text{if } w^Tr \neq 0 \\ 0 & \text{if } w^Tr = 0 \end{cases} . \tag{3.38}$$

When n_s is large enough, $(\tau(w,y))^2$ will be large enough to overcome any given threshold for "significance" for all directions where $w^Tr \neq 0$. This means that if the criterion in equation 3.21 is used to determine n_d, then for n_s sufficiently large $r \in \text{Range}(W_{n_d})$ and $\tilde{b}$ is an unbiased estimator of r.

Beyond the above asymptotic result, one can make other statements about bias. Obviously, $\tilde{b}$ is an unbiased estimate of r whenever $\mathcal{H}_0^{2,n_d+1}$ is true. Moreover, SR strives to choose W_{n_d} so that $r \in \text{Range}(W_{n_d})$, so SR regressors will tend to have the advantages discussed in section 2.3 for restriction regressors when the bias is "small." In fact, empirical work by Cinar [68] and Mejdell [63] has shown that assuming the prediction bias is "small" can be a good assumption for SR, so SR may yield smaller prediction intervals than one would compute using classical methods, as discussed in section 2.3.

To see further benefits of using a restriction regressor directly derived from a statistical foundation, consider the heteroscedastic case, that is $\mathcal{E}\left(ee^T\right) = \sigma_e^2 P$. The SR method begins by computing the minimum-variance unbiased estimator and its variance. Thus, the additional error information is naturally incorporated into the procedure. PLS does *not* make use of this additional information, so PLS is not

equivalent to SR in this case. If one draws draws an analogy to generalized least squares and rescales the data $X_{rescaled} = P^{-\frac{1}{2}}X$ and $y_{rescaled} = P^{-\frac{1}{2}}y$, then performing PLS on the scaled data is equivalent to SR. Thus PLS rests on the assumption of homoscedasticity.

3.4 Why not use the PRESS?

SR was derived directly from the MSE without considering prediction objectives. Thus one might consider deriving a similar regressor explicitly for prediction. For this objective, a reasonable null hypothesis might be

$$\mathcal{H}_0^3 : \qquad < Xr, Xw > = 0. \tag{3.39}$$

For this null hypothesis, a natural test statistic is

$$\tau_{PRESS}(w, y) = \frac{< X\tilde{r}, Xw >}{\sqrt{\text{Var} < X\tilde{r}, Xw >}} = \frac{w^T X^T X \tilde{r}}{\sqrt{w^T X^T X w \sigma_e^2}}. \tag{3.40}$$

Equating the gradient to zero reveals that $w = \tilde{r}$ maximizes equation 3.40; this approach has reproduced the OLS estimator.

A more intriguing question would be to use $\mathcal{H}_0^3$ as the basis for a "prediction significance" test for choosing n_d. Substituting $w_1^{opt} = X^T y$ into $\tau_{PRESS}(w, y)$ under $\mathcal{H}_0^3$ reveals that the resulting test statistic is distributed as $\frac{e^T X X^T e}{\sqrt{e^T X X^T X X^T e}}$. Analyzing such a distribution would lead the current investigation astray and is therefore left as a suggestion for future work.

3.5 Conclusion

This chapter derived the significance regressor, SR, from a null hypothesis, $\mathcal{H}_0^{2,i}$, directly related to improving the MSE of restriction regressors. The popular and successful PLS algorithm was seen to maximize the value of a test statistic for this null hypothesis. Next, a new significance test for evaluating $\mathcal{H}_0^{2,i}$ and determining n_d was developed. As an algebraic formality, one can always reject $\mathcal{H}_0^{2,i}$ with probability one for any non-singular X; see the final developments in the ensuing appendix. However if the *data* cannot muster the strength to reject $\mathcal{H}_0^{2,i}$, which is known to be (formally) false, then one should be wary of using that same data to estimate components of unknown parameters in that subspace. SR was shown to have several useful properties including asymptotic unbiasedness and, in some cases, unbiased prediction. Moreover, SR was derived from a firm statistical viewpoint instead of heuristic motivations, so one can directly generalize the results of this chapter.

3.6 Appendix: Equivalence of SR and PLS for the Scalar Output Case

This appendix links the significance regression method (SR) to PLS for scalar output problems. In particular, a proof is developed that shows Helland's formula for the PLS loading vectors satisfies the necessary condition for the significant subspaces for scalar output models. A "significant vector" is understood to be one of the n_d vectors w_i^{opt} that satisfies equation 3.13. However, also consider the "algebraically-significant vectors." These are the n_A vectors w_i^{opt} that satisfy equation 3.13 for which $\tau(w_i^{opt}, y) \neq 0$.

First let $i-1$ previously determined "significant" vectors be the columns of W_{i-1}^{sig}. Applying the necessary condition of equation 3.13 to the gradient of $(\tau(w,y))^2$, described in equation 3.10, yields the condition

$$(I - W_{i-1}^{sig} W_{i-1}^{sig\,T})(\tilde{r} - \frac{\tilde{\sigma}_\epsilon^2 \left(\tau(w_i^{opt}, y)\right)^2}{\tilde{r}^T w_i^{opt}} (X^T X)^{-1} w_i^{opt}) = 0, \tag{3.41}$$

that must be satisfied in turn by each additional significant vector $w_i^{opt}(y)$.

Next consider Helland's method [29] for computing PLS loading vectors for scalar output problems.

Algorithm 3.2 (Helland's PLS algorithm for scalar output problems)

$$\tilde{r} = (X^T X)^{-1} X^T y \tag{3.42}$$

$$W_0^{pls} = [0|\cdots|0]^T, \quad W_0^{pls} \in \Re^{n_i} \tag{3.43}$$

$$\text{DO} \quad i = 1, n_A$$

$$v = (I - W_{i-1}^{pls} W_{i-1}^{pls\,T})(X^T X)^i \tilde{r} \tag{3.44}$$

$$w_i^{pls} = \frac{v}{\|v\|} \tag{3.45}$$

$$W_i^{pls} = [w_1^{pls}|w_2^{pls}|\cdots|w_i^{pls}] \tag{3.46}$$

END DO.

Now the theorems linking PLS and SR can be stated. Theorem 3.1 shows that the PLS vectors are also "algebraically-significant vectors" when they exist. Additional discussion shows that there are almost surely as many "algebraically-significant vectors" as there are inputs.

Theorem 3.1 *The PLS loading vectors satisfy the necessary condition for all algebraically-significant vectors. That is, if there are n_A algebraically significant vectors, then $W_i^{pls} = W_i^{sig}\ \forall i = 1, \ldots, n_A$. Moreover, any vector satisfying the necessary condition for the ith significant vector is a scalar multiple of w_i^{pls}.*

Proof. The theorem is proven inductively.

For $i = 1$: Substituting the first PLS loading vector $\frac{(X^TX)\tilde{r}}{\|(X^TX)\tilde{r}\|}$ into expression 3.10 yields

$$\nabla_w \left(\tau(w,y)\right)^2 = \frac{2\left(\tilde{r}(r^T(X^TX)\tilde{r})r^T(X^TX)(X^TX)^{-1} - (\tilde{r}^T(X^TX)\tilde{r})^2(X^TX)^{-1}\right)(X^TX)\tilde{r}\tilde{\sigma}_e^2}{(\tilde{r}^T(X^TX)(X^TX)^{-1}(X^TX)\tilde{r}\tilde{\sigma}_e^2)^2} \tag{3.47}$$

$$= \frac{2\tilde{r}(\tilde{r}^T(X^TX)\tilde{r})^2\tilde{\sigma}_e^2 - 2\tilde{r}(\tilde{r}^T(X^TX)\tilde{r})^2\tilde{\sigma}_e^2}{(\tilde{r}^T(X^TX)\tilde{r}\tilde{\sigma}_e^2)^2} \tag{3.48}$$

$$= 0. \tag{3.49}$$

Thus, the first PLS loading vector satisfies the necessary condition for the first significant vector.

Assume true for $i-1$: From algorithm 3.2, the first column of W_{i-1}^{pls} is known to be $\frac{X^TX\tilde{r}}{\|X^TX\tilde{r}\|}$. Moreover, the jth column is $\sum_{k=1}^{j} \alpha_{k,j}(X^TX)^k\tilde{r}$. The scalars $\alpha_{k,j} \neq 0\ \forall\ k \leq j\ \ k < i$ because $(X^TX)^k\tilde{r}$ is linearly independent of $(X^TX)^j\tilde{r}\ \forall j \neq k,\ \ j,k < i$ (due to the existence assumption) and $< (X^TX)^k\tilde{r}, (X^TX)^j\tilde{r} > \neq 0 \quad \forall j,k < i$. The jth column of W_{i-1}^{sig} is also $\sum_{k=1}^{j} \alpha_{k,j}(X^TX)^k\tilde{r}$ because $W_{i-1}^{sig} = W_{i-1}^{pls}$ by assumption.

For i, $i \leq n_A$: From the i-1 step, $W_{i-1} = W_{i-1}^{sig} = W_{i-1}^{pls}$. The necessary condition for the ith significant vector, Equation 3.41, becomes

$$(I - W_{i-1}W_{i-1}^T)\tilde{r} = (I - W_{i-1}W_{i-1}^T)\frac{\tilde{\sigma}_e^2 \tau^2_{w_i^{opt}}}{\tilde{r}^T w_i^{opt}}(X^TX)^{-1}w_i^{opt}. \tag{3.50}$$

Notice that the left-hand side (LHS) of equation 3.50 cannot be zero. If it were, then the ith algebraically-significant vector could not exist.

Now describe w_i^{opt} as the sum of w_i^{pls} and some non-zero vector $v \in$ Range$(I - W_{i-1}W_{i-1}^T)$. This v is distinct from the v in equation 3.44. Then $w_i^{opt} = w_i^{pls} + v$, and equation 3.50 become

$$(I - W_{i-1}W_{i-1}^T)\tilde{r} = (I - W_{i-1}W_{i-1}^T)\frac{\tilde{\sigma}_e^2 \tau^2_{w_i^{opt}}}{\tilde{r}^T w_i^{opt}}(X^TX)^{-1}\left(w_i^{pls} + v\right). \tag{3.51}$$

By extension of the argument from the $i-1$ step, $w_i^{pls} = \sum_{k=1}^{i} \alpha_{k,i}(X^TX)^k\tilde{r}$, where $\alpha_{k,i}$ is still non-zero. Multiplying w_i^{pls} by $(X^TX)^{-1}$ produces $\sum_{k=1}^{i} \alpha_{k,i}(X^TX)^{k-1}\tilde{r}$ which, after multiplying through the projection matrix $(I - W_{i-1}W_{i-1}^T)$ yields $\alpha_{1,i}(I - W_{i-1}W_{i-1}^T)\tilde{r}$. Thus equation 3.51 becomes

$$(I - W_{i-1}W_{i-1}^T)\tilde{r} = \frac{\tilde{\sigma}_e^2 \tau^2_{w_i^{opt}}}{\tilde{r}^T w_i^{opt}}\left(\alpha_{1,i}(I - W_{i-1}W_{i-1}^T)\tilde{r} \; + (I - W_{i-1}W_{i-1}^T)(X^TX)^{-1}v\right). \tag{3.52}$$

Consider now the second term of the RHS. If $\{X^TX\tilde{r}, \ldots, (X^TX)^{n_i}\tilde{r}\}$ spans $\Re^{n_i}$, then one can quickly see that $(I - W_{i-1}W_{i-1}^T)(X^TX)^{-1}v$ is always non-zero. Additionally, one can show that if $\{X^TX\tilde{r}, \ldots, (X^TX)^{n_i}\tilde{r}\}$ does not span $\Re^{n_i}$, then the additional basis vectors needed to span $\Re^{n_i}$ are eigenvectors of X^TX orthogonal to the columns of W_{i-1}. Thus, $(I - W_{i-1}W_{i-1}^T)(X^TX)^{-1}v \neq 0 \; \forall v \in$ Range$(I - W_{i-1}W_{i-1}^T)$.

The LHS of equation 3.52 and the first term of the RHS are both vectors pointing in the same direction, namely $(I - W_{i-1}W_{i-1}^T)\tilde{r}$, so a vector $v \in$ Range$(I - W_{i-1}W_{i-1}^T)$ satisfying $v \neq \eta w_i^{pls}$ for any scalar $\eta \neq 0$ would make equation 3.50 insoluble. However

w_i^{opt} exists by assumption so equation 3.50 must have a well-defined solution and v therefore equals ηw_i^{pls} for some scalar η. Since

- a w_i^{opt} exists that satisfies equation 3.50,
- w_i^{opt} must be a vector pointing in the same direction as w_i^{pls}, and
- equation 3.50 is invariant to the length of w_i^{opt},

$w_i^{opt} = \eta w_i^{pls}$ for any $\eta \neq 0$ must satisfy equation 3.50. □

The PLS loading vectors satisfy the necessary condition for the significant vectors for any $n_d \leq n_A$. Since both PLS and SR compute $\tilde{b}$ using equation 2.8 and the same search space, Range(W_{n_d}), the two methods yield the same $\tilde{b}$. Thus PLS is a useful algorithm for computing the SR search space. However, the above proof raises the question: how many algebraically-significant vectors exist? Drawing directly from Helland's results one knows that n_A (which Helland calls M) is equal to the minimum number of right singular vectors of X (principal component vectors of X) required to form a basis for $\tilde{r}$. See theorem 1 and theorem 2 of Helland [30]. Since $n_A < n_i$ only if $\tilde{r}$ is orthogonal to one of the right singular vectors of X, $n_A = n_i$ almost surely in practice.

Chapter 4

Linear Operator Description of Significance Regression

4.1 Introduction

Significance regression will be generalized in this chapter to encompass the analysis of any linear operator mapping between Hilbert spaces. Two methods are developed: a regression algorithm and a factor analysis algorithm. These algorithms provide a unifying approach for a wide array of models and problems; the SR algorithm in the prior chapter was limited solely to the classical regression model described by equation 2.1. With this description, the results of the prior chapter will be easily generalized to other models and objective functions. This chapter will neither motivate SR nor "prove" any general results; rather, this chapter is descriptive. In doing so, the key features that make significance regression successful are described in a general framework, making the application of SR to other models direct. A brief example concludes this chapter that illustrates the use of significance regression for a problem not normally associated with the classical regression model.

4.2 The General Null Hypothesis

Consider

$$y = R(x) \tag{4.1}$$

where $x \in \mathcal{I}$ is an input, $y \in \mathcal{O}$ is an output, $R \in \mathcal{M}$ is an unknown linear mapping, and $\mathcal{I}$ and $\mathcal{O}$ are Hilbert spaces. $\mathcal{M}$ is the space of linear operators defined by $\mathcal{M} = \mathcal{O} \otimes \mathcal{I}$ where $\otimes$ is the Cartesian (orthogonal) product. Observations of both the input, x, and output, y, may be corrupted by noise. For many problems, n_s observations of (x, y) pairs may be available. Thus, let X and Y be the Cartesian n_s-tuples $X = \{x_1, x_2, \ldots, x_{n_s}\}$ and $Y = \{y_1, y_2, \ldots, y_{n_s}\}$. Notice $X \in \Re^{n_s} \otimes \mathcal{I}$ and $Y \in \Re^{n_s} \otimes \mathcal{O}$. Typically the appropriate inner product for $\Re^{n_s} \otimes \mathcal{O}$ will be the sum over the inner product of samples (*e.g.*, $< Y, Y > = \sum_{i=1}^{n_s} < y_i, y_i >$).

Next, one poses the null hypothesis:

$$\mathcal{H}_0^4 : \qquad < R(\cdot), S(\cdot) > = 0 \tag{4.2}$$

for some $S(\cdot) \in \mathcal{M}$. First, one computes the "best" regressor $\tilde{R} \in \mathcal{M}$, say by minimizing the square of the 2-norm of the error:

$$\tilde{R}(\cdot) = \arg \min_{B(\cdot) \in \mathcal{M}} < Y - \tilde{Y}, Y - \tilde{Y} > \tag{4.3}$$

where $\tilde{Y} = \{B(x_1), B(x_2), \ldots, B(x_{n_s})\}$. A natural test statistic for $\mathcal{H}_0^4$ is

$$\tau(S(\cdot), X, Y) = \frac{< \tilde{R}(\cdot), S(\cdot) >}{\sqrt{\mathrm{Var}(< \tilde{R}(\cdot), S(\cdot) >)}}. \tag{4.4}$$

Notice that τ is invariant to scalar multiplication of $S(\cdot)$: $\tau(\alpha S(\cdot), X, Y) = \tau(S(\cdot), X, Y)$ for any non-zero $\alpha \in \Re$.

4.3 Regression

Next, consider the regression problem. This problem includes both estimating the "true" $R(\cdot)$ and creating a $\tilde{B}(\cdot)$ that is useful for predicting y from observations of x. One proceeds by searching for the "most significant" subspaces of $\mathcal{M}$, represented by the $S(\cdot) \in \mathcal{M}$, that maximize $|\tau(S(\cdot), X, Y)|$. Since

$$\arg \max_{S(\cdot)\in\mathcal{M}} |\tau(S(\cdot), X, Y)| = \arg \max_{S(\cdot)\in\mathcal{M}} (\tau(S(\cdot), X, Y))^2 = S_1^{opt}(\cdot), \tag{4.5}$$

one will usually prefer to work with

$$(\tau(S(\cdot), X, Y))^2 = \frac{(< \tilde{R}(\cdot), S(\cdot) >)^2}{\mathrm{Var}(< \tilde{R}(\cdot), S(\cdot) >)}. \tag{4.6}$$

As noted above, τ is invariant to scalar multiplication of $S(\cdot)$, so $S_1^{opt}(\cdot)$ will not be unique. Assuming the relevant derivatives exist, a necessary condition for the optimum is

$$\nabla_{S(\cdot)} (\tau(S(\cdot), X, Y))^2 \Big|_{S_1^{opt}(\cdot)} = 0. \tag{4.7}$$

When equation 4.7 is meaningful, it yields

$$\nabla_{S(\cdot)} \mathrm{Var}(< S(\cdot), R(\cdot) >) \Big|_{S_1^{opt}(\cdot)} = \frac{2\mathrm{Var}(< S(\cdot), R(\cdot) >)}{< S(\cdot), R(\cdot) >} \nabla_{S(\cdot)} < S(\cdot), R(\cdot) > \Big|_{S_1^{opt}(\cdot)}. \tag{4.8}$$

After one determines S_1^{opt} (whose span is the first "significant subspace"), additional "significant subspaces" are found by repeating the optimization subject to the constraint that all $S_1^{opt}(\cdot), S_2^{opt}(\cdot), \ldots$ must be orthogonal. A "significant subspace" is a subspace spanned by an $S_i^{opt}(\cdot)$. Lastly, the "improved" regressor, $\tilde{B}(\cdot)$, is found via repeating the optimization described in equation 4.3 subject to the constraint the $\tilde{B}(\cdot)$ lie in the space formed from the direct sums of the "significant subspaces." (Alternate statement of the same idea: $\tilde{B}(\cdot)$ is constrained to be a linear combination

of the $S_i^{opt}(\cdot)$). The algorithm to compute n_d significant spaces is

Algorithm 4.1 (General significance regression algorithm)

$$\tilde{R}(\cdot) = \arg \min_{B(\cdot)\in\mathcal{M}} < Y - \tilde{Y}, Y - \tilde{Y} > \tag{4.9}$$

$$\mathcal{S}_0 = \{\ \}, \quad \mathcal{S}_0^{\perp} = \mathcal{M} \tag{4.10}$$

$$\text{DO} \quad i = 1, n_d$$

$$S_i^{opt}(\cdot) = \arg \max_{S(\cdot)\in\mathcal{S}_{i-1}^{\perp}} (\tau(S(\cdot), X, Y))^2 \tag{4.11}$$

$$\mathcal{S}_i = \mathcal{S}_{i-1} \oplus \text{Span}(S_i^{opt}(\cdot)) \tag{4.12}$$

$$\text{END} \quad \text{DO}$$

$$\tilde{B}(\cdot) = \arg \min_{B(\cdot)\in\mathcal{S}_{n_d}} < Y - \tilde{Y}, Y - \tilde{Y} > \tag{4.13}$$

where $\tilde{Y} = \{B(x_1), B(x_2), \ldots, B(x_{n_s})\}$, and $\mathcal{S}_i$, $\mathcal{S}_i^{\perp}$ are complementary Hilbert spaces such that $\mathcal{S}_i \oplus \mathcal{S}_i^{\perp} = \mathcal{M}$ for $i = 0, \ldots, n_d$. Although equation 4.13 is a constrained optimization, the constraints are linear equality constraints so one can typically project the problem into $\mathcal{S}_{n_d}$ and solve an unconstrained problem. Algorithm 4.1 does not reveal how to define n_d; in practice, one has a variety of choices including cross-validation. A useful null hypothesis for the ith significant subspace is

$$\mathcal{H}_0^{5,i}: \ < R(\cdot), S(\cdot) > = 0 \quad \forall S(\cdot) \in \mathcal{S}_{i-1}^{\perp}. \tag{4.14}$$

If one can reject $\mathcal{H}_0^{5,i}$, then $n_d \geq i$.

4.4 Factor Analysis

Often in multivariate analysis one hypothesizes that the "true" structure of a problem can be described in a lower dimensional subspace. Attempts to determine such structure are typically called "factor analysis." Let $w_i \in \mathcal{I}$, $c_i \in \mathcal{O}$, and $t_i \in \Re^{n_s}$.

Then this verbal belief for linear problems can be equivalent to assuming that the X and Y are described by

$$X = t_1 \otimes w_1 + t_2 \otimes w_2 \ \ldots \ t_{n_d} \otimes w_{n_d} + E_X, \tag{4.15}$$

$$Y = \beta_1 t_1 \otimes c_1 + \beta_2 t_2 \otimes c_2 \ \ldots \ \beta_{n_d} t_{n_d} \otimes c_{n_d} + E_Y \tag{4.16}$$

where the β_i are scalars, $E_X \in \Re^{n_s} \otimes \mathcal{I}$, $E_y \in \Re^{n_s} \otimes \mathcal{I}$ are "noise," and $\otimes$ is the appropriate outer product. While the derivations are rigorous, the motivations for the objective functions that drive the various factor analytic methods are notoriously subjective [23] [89, p. 225]; thus no motivation will given here for using SR to compute the "factors" w_i, c_i, and t_i other than to note that these factors will describe "significant" linear relations between the inputs and outputs.

To compute these factors from $(\tau(S(\cdot), X, Y))^2$, one computes the maximum in equation 4.11 subject to the constraint that $S(\cdot)$ be a rank one linear operator. The set of rank one linear operators in $\mathcal{M}$ can be readily parameterized as $\mathcal{F} = \{F(\cdot) \in \mathcal{M} \ | F(\cdot) = c < \cdot, w, > \ \forall w \in \mathcal{I}, c \in \mathcal{O}\}$. Let $b_i \in \Re^i$ be $b_i = [\beta_1, \beta_2, \ldots, \beta_i]^T$. Then the algorithm to compute n_d significant factors is:

Algorithm 4.2 (The General SR Factor Analysis Algorithm)

$$\tilde{R}(\cdot) = \arg \min_{B(\cdot) \in \mathcal{M}} < Y - \tilde{Y}, Y - \tilde{Y} > \tag{4.17}$$

DO $i = 1, n_d$

$$F_i^{opt}(\cdot) = \arg \max_{\substack{F(\cdot) \in \mathcal{F} \text{ s.t.} \\ < F_j(\cdot), F(\cdot) > = 0 \ \forall j < i, \\ < F(\cdot), F(\cdot) > = 1}} (\tau(F(\cdot), X, Y))^2 \tag{4.18}$$

$$t_i = [< w_i^{opt}, x_1 > , < w_i^{opt}, x_2 > , \ldots, < w_i^{opt}, x_{n_s} >]^T \tag{4.19}$$

$$b_i = \arg \min_{b_i \in \Re^i} < Y - \sum_{j=1}^{i} \beta_j F_j^{opt}(X), Y - \sum_{j=1}^{i} \beta_j F_j^{opt}(X) > \tag{4.20}$$

END DO

where $F_j^{opt}(X) = \{F_j^{opt}(x_1), F_j^{opt}(x_2), \ldots, F_j^{opt}(x_{n_s})\}$ is a Cartesian n_s-tuple. This particular algorithm states the unnecessary constraint $< F(\cdot), F(\cdot) > = 1$. However, since one usually performs factor analysis intending to "interpret" the factors, factors are typically normalized. The user is free to choose a different normalization than the one used here.

Some factor analysis choose n_d by testing E_x and/or E_y against the null hypothesis that these objects are only noise. One can certainly adapt these method to the SR factor analysis approach. Another natural null hypothesis to help determine n_d is

$$\mathcal{H}_0^{6,i}: \quad < R(\cdot), F(\cdot) > = 0 \quad \forall F(\cdot) \in \mathcal{F} \text{ s.t. } < F_j(\cdot), F(\cdot) > = 0 \ \ \forall j < i. \quad (4.21)$$

The special constraints on the search space of the factor analysis problem allow one to state additional null hypothesis. Define the spaces $\mathcal{W}_i^{\perp} = \{w \in \mathcal{I} | < w, w_j^{opt} > = 0 \ \forall j < i\}$ and $\mathcal{C}_i^{\perp} = \{c \in \mathcal{O} | < c, c_j^{opt} > = 0 \ \forall j < i\}$. Using these spaces,

$$\mathcal{H}_0^{7,i}: \quad \langle R(\cdot) \ , \ c < w, \cdot > \rangle = 0 \quad \forall w \in \mathcal{W}_i^{\perp}, \quad (4.22)$$

$$\mathcal{H}_0^{8,i}: \quad \langle R(\cdot) \ , \ c < w, \cdot > \rangle = 0 \quad \forall c \in \mathcal{C}_i^{\perp}, \text{ and} \quad (4.23)$$

$$\mathcal{H}_0^{9,i}: \qquad R(x) = 0 \qquad \forall x \in \mathcal{W}_i^{\perp} \quad (4.24)$$

may be computationally simpler. $\mathcal{H}_0^{7,i}$ and $\mathcal{H}_0^{8,i}$ follow directly from $F(\cdot) = c < \cdot, w_i >$. If $\mathcal{H}_0^{9,i}$ holds then

$$< R(x), c < x, w_i > > = 0 \ \ \forall x \in \mathcal{W}_i^{\perp}. \quad (4.25)$$

For any x in the orthogonal complement of $\mathcal{W}_i^{\perp}$, expression 4.25 is also zero because $w_i \in \mathcal{W}_i^{\perp}$. Therefore $\mathcal{H}_0^{9,i} \Rightarrow \mathcal{H}_0^{6,i}$. All three null hypotheses $\mathcal{H}_0^{7,i}$, $\mathcal{H}_0^{8,i}$ and $\mathcal{H}_0^{9,i}$ are sufficient conditions for $\mathcal{H}_0^{6,i}$; rejecting these does not necessarily reject $\mathcal{H}_0^{6,i}$. Still, for any given problem, computational tractability may make these sufficient conditions

more practical than the logically rigorous $\mathcal{H}_0^{6,i}$.

4.5 A System Identification Illustration

The remainder of this thesis works with $\mathcal{I}$ and $\mathcal{O}$ that are finite dimensional vector spaces; the results of this chapter are not essential for the ensuing development. However, this chapter is useful for translating the results of this thesis to problems beyond the usual statistical regression problems. This section presents a brief (and naive) example of using significance regression for system identification. No claim is made that the SR system identification method developed here is useful; investigating such a claim would carry this study far beyond its proper scope. The example does succinctly demonstrate the implementation of SR in a new area via the operator theoretic description of this chapter.

The notation used here will deviate from the standard notation of this thesis so as to conform to Ljung's notation [56]. Consider a linear system g that relates the input $u(t)$ to the output $y(t)$ via the convolution

$$y(t) = \sum_{k=1}^{\infty} g(t)u(t-k) + \nu(t) \quad t = 0, 1, 2, \ldots \tag{4.26}$$

where $\nu(t)$ is a zero-mean noise process with spectrum $\Phi_\nu(\omega)$; the spectrum is defined in Ljung [56]. The discrete Fourier transform (DFT) of the input and output is

$$U_N(\omega) = \frac{1}{\sqrt{N}} \sum_{t=1}^{N} u(t)e^{-i\omega t} \text{ and} \tag{4.27}$$

$$Y_N(\omega) = \frac{1}{\sqrt{N}} \sum_{t=1}^{N} y(t)e^{-i\omega t} \tag{4.28}$$

where $\omega = 2\pi k/N,\ k = 1, 2, \ldots, N$. Recalling that the Fourier basis is orthonormal, define the inner product to be $< Y_N(\omega), U_N(\omega) > = \sum_\omega Y_N(\omega)U_N(\omega)$. Estimate the

DFT of $g(t)$ as

$$\tilde{G}_N(\omega) = \arg\min_{G_N(\omega)} < Y_N(\omega) - G_N(\omega)U_N(\omega), Y_N(\omega) - G_N(\omega)U_N(\omega) > \quad (4.29)$$

$$= \frac{Y_N(\omega)}{U_N(\omega)}. \quad (4.30)$$

The minimization is performed by imposing the following ordering on the complex numbers: for any two complex numbers a and b, $a > b$ if $|a| > |b|$ or if $|a| = |b|$ and $\arg(a) > \arg(b)$. $\tilde{G}_N(\omega)$ is known as the empirical transfer-function estimate (ETFE). If $u(t)$ is a realization of a stochastic process, then the ETFE is an asymptotically unbiased estimate of the transfer function and the estimates at different frequencies are asymptotically uncorrelated. [56, p. 150].

$\tilde{G}_N(\omega)$ is known to be noise sensitive. Typically, the ETFE is "smoothed" by assuming that the "true" transfer function is a smooth function of ω. Such an approach, also known as regularization, is explicitly Bayesian and can be interpreted as relating to the ridge regression approach of section 2.1. SR provides an alternative method. Consider a new function $W_N(\omega)$ and the null hypothesis

$$\mathcal{H}_0^{10}: \qquad < G_N(\omega), W_N(\omega) > = 0. \quad (4.31)$$

The variance needed for the test statistic is

$$\text{Var}\left(< \tilde{G}_N(\omega), W_N(\omega) > \right) = \sum_{\omega=1}^{N} \left(W_N(\omega)\right)^2 \Phi_\nu(\omega). \quad (4.32)$$

From here, one can use algorithm 4.1 to compute the SR transfer function estimate.

The "smoothing" methods have a long history linked to a strong and pervasive engineering heuristic, so believing that the simplistic SR method presented here represents a step forward for system identification methods is optimistic. However, problems in which the "smoothing" assumption is not well motivated may benefit from an

SR-based method. An example is the identification of the resonant modes in lightly damped structures. For these structures the Bode plots are known to be "spikey"; the smoothness assumption still holds, but in a weaker sense. For control purposes, one is often interested in the location and magnitude of these modes; since "smoothing" will tend to "knock down" these spikes, the SR approach may be more appealing.

4.6 Summary

This chapter presented a linear operator theoretic description of significance regression. The null hypothesis that drives the method was formulated for linear operators mapping between Hilbert spaces. This null hypothesis was then used as the basis for a biased regression algorithm and a factor analysis method. The additional constraints imposed by the factor analysis method led to the formulation of additional null hypothesis sufficient for the general null hypothesis. The utility of the the general description was illustrated by using the linear operator framework to generate a (naive) system identification method.

Part III

Generalizations and Improvements for Significance Regression

Chapter 5

Multiple Output Significance Regression

5.1 Introduction

The methods of chapter 3 dealt only with scalar output problems. This chapter uses the linear operator description of the prior chapter to extend significance regression to problems with multiple outputs. Sections 5.2, 5.3 and 5.4 work with vector output problems of the form

$$Y = XR + E, \tag{5.1}$$

where $Y \in \Re^{n_s \times n_o}$ is known, $R \in \Re^{n_i \times n_o}$ is an unknown regression matrix, and $E \in \Re^{n_s \times n_o}$ is an unobservable matrix of errors. For simplicity of development, further assume that the elements of E are zero-mean, independent, and homoscedastic random variables: $\mathcal{E}(E) = 0$, $\mathcal{E}\left(E^T E\right) = n_s \sigma_e^2 I$ and $\mathcal{E}\left(EE^T\right) = n_o \sigma_e^2 I$. The independence and homoscedasticity assumptions can be readily relaxed; see section 8.5 for further discussion. First the vector output regression algorithm and factor analysis algorithm are developed. Next, the relationship between PLS and SR is analyzed, followed by a brief discussion of the SR approach to tensor problems.

For the input space $\mathcal{I} = \Re^{n_i}$ and output space $\mathcal{O} = \Re^{n_o}$, the natural inner product for $\mathcal{M} = \Re^{n_o \times n_i}$, the space of linear operators mapping from $\mathcal{I}$ to $\mathcal{O}$, is the tensor inner product. This inner product is defined as $< A, B > = \mathrm{Tr}(AB^T)$ and is the inner product that defines the matrix Frobenius norm: $\sqrt{< A, A >} = ||A||_F$.

Following the description of section 4.3, the null hypothesis of interest is

$$\mathcal{H}_0^{11} : \qquad < R, S > = 0 \tag{5.2}$$

where $S \in \Re^{n_i \times n_0}$. For this null hypothesis, a natural test statistic is

$$\tau(S, Y) = \frac{< \tilde{R}, S >}{\sqrt{\mathrm{Var} < \tilde{R}, S >}} \tag{5.3}$$

$$= \frac{\mathrm{Tr}(\tilde{R} S^T)}{\sqrt{\mathrm{Tr}\left(S^T (X^T X)^{-1} S\right) \sigma_e^2}}. \tag{5.4}$$

With these definitions one can directly develop SR algorithms for both regression and factor analysis.

5.2 Regression

One can develop SR directly from $\tau(S, Y)$ using tensor operations. However, a much simpler approach is available. By noting that $\Re^{n_o \times n_i}$ is isomorphic and isometric to $\Re^{n_o n_i}$ one can recast the vector output problem as a scalar output problem and use the results of chapter 3. To use the scalar output algorithms, the input data matrix and the output data matrix need to be suitably redefined. Normally,

$$Y = \begin{bmatrix} y_1^T \\ \vdots \\ y_{n_s}^T \end{bmatrix}, \qquad X = \begin{bmatrix} x_1^T \\ \vdots \\ x_{n_s}^T \end{bmatrix}, \quad \text{and } Y = XR + E. \tag{5.5}$$

To conform to equation 2.1, Y, E, and R must be transformed into column vectors. Considering the columns of $R = [r_1| \cdots |r_{n_o}]$, let

$$y_{stacked} = \begin{bmatrix} y_1 \\ \vdots \\ y_{n_s} \end{bmatrix} \qquad \text{and } r_{stacked} = \begin{bmatrix} r_1 \\ \vdots \\ r_{n_o} \end{bmatrix} \tag{5.6}$$

where $y_{stacked} \in \Re^{n_s n_o}$ and $r_{stacked} \in \Re^{n_i n_o}$. Create $e_{stacked}$ from E is the same manner that $y_{stacked}$ was created from Y. Moreover, build $X_{stacked} \in \Re^{n_s n_o \times n_i n_o}$ such that

$$X_{stacked} = \overbrace{\begin{bmatrix} \left.\begin{matrix} x_1^T & 0 & \dots & 0 \\ 0 & x_1^T & & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \dots & x_1^T \end{matrix}\right\} n_o \\ \begin{matrix} x_2^T & 0 & \dots & 0 \\ 0 & x_2^T & & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \dots & x_2^T \end{matrix} \\ \vdots \\ \begin{matrix} x_{n_s}^T & 0 & \dots & 0 \\ 0 & x_{n_s}^T & & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \dots & x_{n_s}^T \end{matrix} \end{bmatrix}}^{n_i \times n_o}. \tag{5.7}$$

Then equation 5.1 data can be described by

$$y_{stacked} = X_{stacked}\ r_{stacked} + e_{stacked}. \tag{5.8}$$

Equation 5.8 is consistent with equation 2.1, so algorithm 3.1 can be used. After the SR process is completed, $\tilde{b}_{stacked}$ must be "unstacked" in the reverse manner from which $r_{stacked}$ was built in equation 5.6.

5.3 Factor Analysis

As described in chapter 4, a common objective of multivariable analysis is to develop a lower dimensional description of data in which the "useful" information has been preserved. For example, principal components analysis has long been used to identify a small set of "loading vectors" that encompass the greatest portion of the variance of a set of data. More recently, several practitioners have recommended PLS when one is interested in variances of input data that explain variances in the output variables. [20, 62]. The approach has been found particularly effective for multivariable stochastic process control [51, 71]. To address this desire, the SR factor analysis algorithm for vector outputs is now developed.

To evaluate vectors $w \in \Re^{n_i}$ and $c \in \Re^{n_0}$ using $\tau(S,Y)$, one parameterizes the rank one matrix (linear operator) as $S = wc^T$. Equation 5.4 then becomes

$$\tau(w,c,y) = \frac{w^T \tilde{R} c}{\sqrt{w^T (X^T X)^{-1} w \; c^T c \; \sigma_e^2}} \tag{5.9}$$

which, if normal errors are assumed, arises from a normal distribution for any given w and c. Next, one solves for the optimal w and c. The wc^T parameterization yields $\nabla_S(\cdot) = \begin{bmatrix} \nabla_w(\cdot) \\ \nabla_c(\cdot) \end{bmatrix}$. Solving

$$\nabla_w \; \tau^2(w,c,Y)\Big|_{S^{opt}(Y)} = 0 \tag{5.10}$$

for $w^{opt}(y)$ one finds

$$w^{opt} = X^T X \tilde{R} c^{opt}(y) = X^T Y c^{opt}(y). \tag{5.11}$$

Solving

$$\nabla_c \ (\tau(w, c, Y))^2 \Big|_{S^{opt}(Y)} = 0 \tag{5.12}$$

and dropping the arguments results in

$$\tilde{R}^T w^{opt} = \frac{w^{opt^T} \tilde{R} c^{opt}}{c^{opt^T} c^{opt}} c^{opt}. \tag{5.13}$$

Substituting w^{opt} from equation 5.11 one finds that c^{opt} satisfies

$$\tilde{R}^T X^T X \tilde{R} c^{opt} = \sigma_e^2 \ \tau_4^2 \Big|_{S^{opt}(Y)} c^{opt}. \tag{5.14}$$

Thus finding the eigenvector of maximum eigenvalue in equation 5.14 yields the c^{opt} needed for equation 5.11. Since $\mathcal{E}\left(\tilde{R}^T X^T X \tilde{R}\right)$ and $\mathcal{E}\left(Y^T Y\right)$ have the same eigenvectors when the data conform to equation 5.1 one can see that in using the "most significant subspace" one is selecting the c that explains the greatest variance in the output data. This is in contrast to PLS, where c is chosen as the eigenvector of $Y^T X X^T Y$ with maximum eigenvalue.

To find additional "significant vectors," one repeats the process enforcing the orthogonality constraint $< S_i^{opt}(Y), S_j^{opt}(Y) > = 0$. Due the to wc^T parameterization, $< S_i^{opt}(Y), S_j^{opt}(Y) > = 0$ if $< w_i^{opt}(Y), w_j^{opt}(Y) > = 0$ or $< c_i^{opt}(Y), c_j^{opt}(Y) > = 0$. The method developed below relies on the two assumptions: $n_i \geq n_o$ and $< w_i^{opt}(Y), w_j^{opt}(Y) > = 0 \ \forall i \neq j$. An equivalent method for the case $n_i < n_o$ can be developed by invoking the alternate second assumption $< c_i^{opt}(Y), c_j^{opt}(Y) > = 0 \ \forall i \neq j$.

Any matrix $W_{i-1} \in \Re^{n_i \times n_w}$ can be factored as $W_{i-1} = Q_{\mathrm{QR}} R_{\mathrm{QR}}$ where $Q_{\mathrm{QR}} \in$

$\Re^{n_i \times n_i}$ is an orthogonal matrix and $R_{\mathrm{QR}} \in \Re^{n_i \times n_w}$ is an upper triangular matrix. A property of the QR decomposition is that if one partitions $Q_{\mathrm{QR}} = [Q_{\mathrm{QR},1}, Q_{\mathrm{QR},2}]$ such that $Q_{\mathrm{QR},1}$ contains the first n_w columns of Q_{QR}, then the columns of $Q_{\mathrm{QR},2}$ describe the null space of W_{i-1}^T. Let the columns of W_{i-1} be $w_1^{opt}(Y)$ through $w_{i-1}^{opt}(Y)$ and $W_{i-1}^{\perp} = Q_{\mathrm{QR},2}$. For this problem the necessary condition for a constrained extremum is

$$W_{i-1}^T \ \nabla_w \ (\tau(w,c,Y))^2 \Big|_{S^{opt}(Y)} = 0, \tag{5.15}$$

which gives rise to the eigenvector equation

$$\sigma_e^2 \tau^2 q = {W_{i-1}^{\perp}}^T X^T X \tilde{R} \tilde{R}^T W_{i-1}^{\perp} \ q \tag{5.16}$$

where $q \in \Re^{n_p}$ and the arguments have been suppressed. The q which is the eigenvector of maximum eigenvalue in equation 5.16 yields $w_i^{opt}(Y) = W_{i-1}^{\perp} q$. The full algorithm is

Algorithm 5.1 (SR-Factor. Factor analysis for vector outputs)

$$\tilde{R} = (X^T X)^{-1} X^T Y \tag{5.17}$$

$$W_0 = [0 \ \ldots \ 0]^T, \quad W_0 \in \Re^{n_i} \tag{5.18}$$

DO i = 1 to n_d

Perform QR factorization of W_{i-1}

$$W_{i-1} \Rightarrow Q_{\mathrm{QR}}, \ R_{\mathrm{QR}} \tag{5.19}$$

$$W_{i-1}^{\perp} = \text{last } n_p = n_i - i + 1 \text{ columns of } Q_{\mathrm{QR}} \tag{5.20}$$

q = eigenvector of maximum eigenvalue of

$${W_{i-1}^{\perp}}^T X^T X \tilde{R} \tilde{R}^T W_{i-1}^{\perp} \tag{5.21}$$

$$w_i^{opt}(Y) = W_{i-1}^{\perp} q \ / ||W_{i-1}^{\perp} \ q|| \tag{5.22}$$

$$c_i^{opt}(Y) = \tilde{R}^T w_i^{opt} \tag{5.23}$$

$$W_i \quad = [w_1^{opt}|w_2^{opt}|\cdots|w_i^{opt}] \tag{5.24}$$

END DO.

5.4 Relationship to Partial Least Squares

As noted above, a currently successful algorithm for factor analysis is partial least squares (PLS); a review of PLS is in appendix A. The factors derived from PLS have also been used successfully for defining restriction regressors. Moreover PLS is closely related to SR; the only difference between algorithm 5.1 and PLS is in the specification of c_i^{opt}. Equation 5.11, which determines w_i^{opt}, corresponds to the formula for PLS for determining w_i^{opt} (see equation A.5).

SR chooses vectors that explain the greatest variance in Y^TY when equation 5.1 describes the data. To see this, use the alternate but equivalent second assumption $< c_i^{opt}(Y), c_j^{opt}(Y) > = 0 \ \forall i \neq j$. Under this assumption, the condition for a constrained extremum gives rise to the eigenvector equation

$$\sigma_e^2\tau^2 q = {C_{i-1}^{\perp}}^T \tilde{R}^T X^T X \tilde{R} C_{i-1}^{\perp} \ q \tag{5.25}$$

where $C_{i-1}^{\perp}$ spans the space of allowable $c_i^{opt}(Y)$. Since

$$\mathcal{E}\left({C_{i-1}^{\perp}}^T \tilde{R}^T X^T X \tilde{R} C_{i-1}^{\perp}\right) = \mathcal{E}\left({C_{i-1}^{\perp}}^T Y^T Y C_{i-1}^{\perp}\right), \tag{5.26}$$

SR computes the vector in Range($C_{i-1}^{\perp}$) that explains the greatest variance in Y^TY. PLS chooses vectors using the slightly different criterion $Y^TXX^TY = \tilde{R}^T(X^TX)^2\tilde{R}$, so PLS can be viewed as a very close approximation to the SR method for building "significant factors." Less formally, the difference between SR and PLS is that SR adheres strictly to the model $Y = XR + E$ and assumes any variation in Y is either a linear function of the variance in X or zero-mean noise. PLS works under the

assumption that there may be "factors" of Y that are not linear functions of X [35].

The difference between the SR method for regression and PLS is more substantial than for factor analysis and can be easily seen by considering the S used in the two approaches. In PLS, S is constrained to only those matrices that can be described wc^T; thus, S is rank one and has only $n_o + n_i$ free parameters. However in SR regression, S is allowed to be any matrix, and thus has $n_o \times n_i$ free parameters. Generally, S_i^{opt} will be full rank (after all, $S_1^{opt} = \alpha X^T Y$). However to describe a full rank S from the PLS vectors, one must build at least n_o PLS directions, assuming $n_i \geq n_o$. Building a full-rank S_{PLS}^{opt} from the first n_o PLS loading vectors leads to $\left(\tau(S_{PLS}^{opt}, Y)\right)^2 \geq \left(\tau(S_1^{opt}, Y)\right)^2$; PLS is sub-optimal for determining the "most significant subspace." Another PLS variant, "one-at-a-time" PLS (OAT-PLS) is closely related to SR. In OAT-PLS one ignores the vector output problem and solves n_o "independent" single-output problems. However, OAT-PLS is also sub-optimal for determining the "most significant subspace." See [35] for further discussion of OAT-PLS and also the properties of PLS for vector output problems.

5.5 Tensor Problems

Tensors, which are sometimes referred to as "multi-way arrays" or "matrices of matrices," are generalizations of matrices that allow one to work with a richer class of problems than is encompassed by multivariable regression. As improvements in computer technology have made larger and richer data structures more readily available to practitioners, more researchers have been pondering tensor data descriptions [77, 99]; tensors are particularly interesting for problems where multiple sensors are being used [10].

For tensors an appropriate inner product can still be described by $< A, B >$ $= \text{Tr}(AB^T)$. That is, for any tensors $A, B \in \Re^{n_1 \times \ldots \times n_N}$, $< A, B >$ $=$

$\sum_{i_1,\ldots,i_M} a_{i_1,\ldots,i_M} b_{i_1,\ldots,i_M}$. If one desired to build a tensor restriction regressor, then the tensors can be "stacked" and one can use algorithm 3.1 and the results of Chapter 3. If one desired to build "significant factors," then one would define a tensor S from the outer product of a vector from each of the component vector spaces and optimize equation $\tau(S, Y)$. Such an approach generates "power-law" algorithms familiar to PLS researchers.

5.6 Summary

The application of the SR method to multiple output problems was carried out. After defining the necessary inner product and null hypothesis, the vector output regression problem was seen to be equivalent to the scalar output problem; by properly "stacking" the vector output problem one can use algorithm 3.1 and the results of chapter 3. Because the vector output problem can be so easily reduced to the scalar problem, only the scalar problem will be considered in the remaining chapters. The SR factor analysis method for vector outputs was shown to be almost identical to the conspicuously successful PLS algorithm. PLS's deficiencies for regression were highlighted via this comparison. Lastly, the issue of tensor data problems was briefly examined. For regression, the tensor problem also "stacks" to the scalar problem.

Chapter 6

Measurement Error Model Significance Regression

6.1 Introduction

The classical model (equation 2.1) assumes that the inputs (the explanatory variables) are known without error. However practitioners must often work with data where all the measurements are corrupted by measurement noise, not just the variables than one wishes to regress onto (the dependent variables). This chapter addresses regression problems where the data are assumed to be described by the measurement error model (MEM)

$$y = Tr + e, \quad \text{and} \tag{6.1}$$

$$X = T + E. \tag{6.2}$$

In this formulation, $T \in \Re^{n_s \times n_i}$ represents the "true" but unobservable explanatory variables, while $X \in \Re^{n_s \times n_i}$ and $y \in \Re^{n_s}$ represent the n_s observations of the explanatory and dependent variables, respectively. Since this work focuses on collinear

problems, the main results will be most applicable to problems where the condition number [21] of $T^T T$ is "large." The unobservable errors affecting the explanatory and dependent variables are $E \in \Re^{n_s \times n_1}$ and $e \in \Re^{n_s}$, respectively. The assumptions used in this treatment are:

(A1) E and T are stochastically independent.

(A2) The elements of T are fixed (but unknown) variates.

(A3) $\mathcal{E}(E) = 0$, $\mathcal{E}(e) = 0$.

(A4) E and e are stochastically independent.

(A5) The fourth moments of all the components of E and e exist.

(A6) $\lim_{n_s \to \infty} \frac{1}{n_s} T^T T = M_T$ exists and is non-singular.

(A7) Each row of E, e_i^T, is stochastically independent and identically distributed, with $\mathcal{E}\left(e_i e_i^T\right) = \Sigma$ and Σ non-singular. Likewise, the elements of e are also stochastically independent and homoscedastic with $\mathcal{E}\left(ee^T\right) = \sigma_e^2 I$.

(A8) In addition to (A6), as $n_s \to \infty$, $T^T T \to n_s M_T$.

Further define $M_X = M_T + \Sigma$. (A4) can be readily relaxed, but at the expense of more involved notation. (A2) can be relaxed by assuming that all limits are regular with probability one. For this more general assumption, all of the results of this chapter hold "conditioned on T." See Schneeweiß [83] for further discussion of these assumptions, their implications, and how to relax them.

For this model an asymptotically unbiased estimate of r is

$$\tilde{r}_{\text{MEM}} = (X^T X - n_s \Sigma)^{-1} X^T y. \tag{6.3}$$

Under the additional assumption of normally distributed errors $\tilde{r}_{\text{MEM}}$ is the maximum likelihood estimate of r [44, 55]. Schneeweiß [83] has also determined the asymptotic properties of $\tilde{r}_{\text{MEM}}$. Specifically, $\sqrt{n_s}(\tilde{r}_{\text{MEM}} - r)$ has an asymptotically normal distribution with variance

$$\text{Var}_\infty\left(\sqrt{n_s}(\tilde{r}_{\text{MEM}} - r)\right) = \lim_{n_s \to \infty} n_s(\tilde{r}_{\text{MEM}} - r)(\tilde{r}_{\text{MEM}} - r)^T \tag{6.4}$$

$$= M_T^{-1}(\Sigma r r^T \Sigma + \sigma_e^2 M_X) M_T^{-1}. \tag{6.5}$$

Consider for the moment the "no input noise" case, $\Sigma = 0$. Then $\text{Var}_\infty\left(\sqrt{n_s}(\tilde{r}_{\text{MEM}} - r)\right) = \sigma_e^2 M_T^{-1}$. If collinearities exist among the explanatory variables then M_T will be "nearly" singular and the variance will tend to be large; this effect is well-known for classical least-squares regressors. The input noise adds additional variance ($M_T^{-1}\Sigma r r^T \Sigma M_T^{-1}$) while maintaining vulnerability to ill-conditioned data ($\sigma_e^2 M_T^{-1} M_X M_T^{-1}$), so clearly one needs to be able to address this problem for the MEM as well.

6.2 Significance Regression Method

For treating collinearity via significance regression one considers the null hypothesis

$$\mathcal{H}_0^{12}: \qquad < w, r > = 0 \tag{6.6}$$

for which a natural test statistic is

$$\tau_{ideal}\ (w, X, y) = \frac{< w, \tilde{r}_{\text{MEM}} >}{\sqrt{\text{Var}(< w, \tilde{r}_{\text{MEM}} >)}}. \tag{6.7}$$

Computing $\mathrm{Var}(< w, \tilde{r}_{\mathrm{MEM}} >)$ is involved; however, one can use equation 6.5 to discern that as $n_s \to \infty$

$$\mathrm{Var}(< w, \tilde{r}_{\mathrm{MEM}} >) \to \frac{1}{n_s} w^T M_T^{-1} (\Sigma r r^T \Sigma + \sigma_e^2 M_X) M_T^{-1} w = V_{ideal}. \tag{6.8}$$

Thus, in principle one can use a t-test based on equation 6.7 to evaluate $\mathcal{H}_0^{12}$ for any given w when n_s is large. However, V_{ideal} includes several terms that must themselves be estimated. First, one must determine M_T. By (A2) the T are fixed variates. One may therefore assume that the values of T are repeated as $n_s \to \infty$, consider M_T to be $\frac{1}{n_s} T^T T$, and approximate M_X as $\frac{1}{n_s} X^T X$. A more difficult problem is the explicit appearance of the unknown vector r in equation 6.8. A useful choice is to use the MLE estimate $\tilde{r}_{\mathrm{MEM}}$ in place of r, acknowledging that this alters the distributional properties of $\tau(w, X, y)$. This substitution has not had major consequence in the various simulation studies conducted thus far; see section 9.3 for further illustration.

With these approximations, the new approximate test statistic is

$$\tau(w, X, y) = \frac{< w, \tilde{r}_{\mathrm{MEM}} >}{\sqrt{\frac{1}{n_s} w^T (\frac{X^T X}{n_s} - \Sigma)^{-1} (\Sigma \tilde{r}_{\mathrm{MEM}} \tilde{r}_{\mathrm{MEM}}^T \Sigma + \sigma_e^2 \frac{X^T X}{n_s}) (\frac{X^T X}{n_s} - \Sigma)^{-1} w}}. \tag{6.9}$$

Applying algorithm 4.1 to the measurement error model using the quantities defined here, the significance regression algorithm for measurement error models is:

Algorithm 6.1 (SR-MEM.)

$$\tilde{r}_{\mathrm{MEM}} = (X^T X - n_s \Sigma)^{-1} X^T y \tag{6.10}$$

$$V = \frac{1}{n_s} \left(\frac{X^T X}{n_s} - \Sigma \right)^{-1} \left(\Sigma \tilde{r}_{\mathrm{MEM}} \tilde{r}_{\mathrm{MEM}}^T \Sigma + \sigma_e^2 \frac{X^T X}{n_s} \right) \left(\frac{X^T X}{n_s} - \Sigma \right)^{-1} \tag{6.11}$$

$$W_0 = [0 \cdots 0]^T, \quad W_0 \in \Re^{n_i} \tag{6.12}$$

DO $i = 1, n_d$

$$w_i^{opt}(X,y) = \frac{(I - W_{i-1}W_{i-1}^T)V^{-i}\tilde{r}_{\text{MEM}}}{||(I - W_{i-1}W_{i-1}^T)V^{-i}\tilde{r}_{\text{MEM}}||} \tag{6.13}$$

$$W_i = [w_1^{opt}|w_2^{opt}|\cdots|w_i^{opt}] \tag{6.14}$$

END DO

$$\tilde{b} = W_{n_d}(W_{n_d}^T(X^TX - n_s\Sigma)W_{n_d})^{-1}W_{n_d}^TX^Ty. \tag{6.15}$$

The w_i^{opt} and the associated $(\tau_i^{opt}(X,y))^2 = (\tau(w_i^{opt},X,y))^2$ have several useful properties. First, assume w_i^{opt} was computed using V_{ideal}. Then

$$(\tau_{ideal,1}^{opt}(X,y))^2 = \frac{(\tilde{r}_{\text{MEM}}^T w_1^{opt})^2}{w_i^{opt^T} V_{ideal} w_1^{opt}} \tag{6.16}$$

$$= \frac{(\tilde{r}_{\text{MEM}}^T V_{ideal}^{-1}\tilde{r}_{\text{MEM}})^2}{\tilde{r}_{\text{MEM}}^T V_{ideal}^{-1} V_{ideal} V_{ideal}^{-1}\tilde{r}_{\text{MEM}}} \tag{6.17}$$

$$= \tilde{r}_{\text{MEM}}^T V_{ideal}^{-1}\tilde{r}_{\text{MEM}} \tag{6.18}$$

which asymptotically approaches a non-central χ^2-distribution with n_i degrees of freedom; when $\mathcal{H}_0^{12}$ holds for w_1^{opt}, then $(\tau_{ideal,1}^{opt}(X,y))^2$ approaches a central χ^2-distribution. Based on this observation, one can approximate the distribution of $(\tau_1^{opt}(X,y))^2$ as being a non-central χ^2-distribution with n_i degrees of freedom. As developed in chapter 3, one can likewise approximate the distribution of $(\tau_i^{opt}(X,y))^2$ with a non-central χ^2-distribution with $n_p = n_i - i + 1$ degrees of freedom. This approximation will breakdown when the value of the non-centrality parameter approaches n_p. Importantly, SR seeks the directions for which the non-centrality parameter dominates the variance. When the approximation fails, the non-central χ^2-distribution will have too heavy of a tail relative to the "true" distribution.

These distributions are used to determine n_d. For each i, one should test if there are any more "significant subspaces" to be found. If no such subspaces exist, then

one should choose $n_d = i$. This motivates the null hypothesis

$$\mathcal{H}_0^{13,i}: \qquad < w, r > = 0 \quad \text{for all } w \in \text{Range}(I - W_{i-1}W_{i-1}^T). \tag{6.19}$$

Under $\mathcal{H}_0^{13,i}$, the non-centrality parameter of the distribution for $(\tau_i^{opt}((X,y))^2$ is zero, so one may use a χ^2 test for determining n_d. If one can reject $\mathcal{H}_0^{13,i}$, then $n_d > i$. The n_d determined using the approximate χ^2 test described here will be less than or equal to the n_d one would determine using the "true" distribution since the approximate distribution has too heavy of a tail.

Although $\tilde{b}$ will typically provide a biased estimate for r, $\tilde{b}$ is in fact asymptotically unbiased. For any w,

$$\lim_{n_s \to \infty} (\tau(w,y))^2 = \lim_{n_s \to \infty} \frac{(w^T \tilde{r}_{\text{MEM}})^2}{w^T V w} \tag{6.20}$$

$$= \lim_{n_s \to \infty} \frac{n_s (w^T \tilde{r}_{\text{MEM}})^2}{n_s w^T V w} \tag{6.21}$$

$$= \begin{cases} \infty & \text{if } w^T r \neq 0 \\ 0 & \text{if } w^T r = 0 \end{cases}. \tag{6.22}$$

When n_s is large enough, $(\tau(w,y))^2$ will be large enough to overcome any given threshold for "significance" for all directions where $w^T r \neq 0$. This means that if the χ^2 test described above is used to determine n_d, then for n_s sufficiently large $r \in \text{Range}(W_{n_d})$ and $\tilde{b}$ is an asymptotically unbiased estimator of r.

6.3 Classical Regression Model Methods

One does not always know the value of Σ. However, this section will show that one can still successfully use techniques derived for the classical model, equation 2.1, for collinear data described by the measurement error model. Using the classical ordinary least-squares (OLS) regressor $\tilde{r}$ for a measurement error model produces an

asymptotically biased estimate of r since $\lim_{n_s \to \infty}(\tilde{r} - r) = -M_X^{-1}\Sigma r \neq 0$. However, $\tilde{r}$ is an asymptotically unbiased estimate of the least-squares optimal predictor p; see claim 6.1 and claim 6.2 of section 6.4, Berkson [1], and Schneeweiß[83] for further discussion. Moreover, $\tilde{r}$ acts as a "natural ridge regressor." $\tilde{r}_{\text{MEM}}$ can be described by

$$\tilde{r}_{\text{MEM}} = \arg \min_{v \in \Re^{n_i}} (y - Xv)^T (y - Xv) - n_s v^T \Sigma v. \tag{6.23}$$

To form an MEM ridge regressor, add the "soft" constraint

$$\tilde{r}_{\text{MEM}} = \arg \min_{v \in \Re^{n_i}} (y - Xv)^T (y - Xv) - n_s v^T \Sigma v + v^T A v \tag{6.24}$$

for any positive semi-definite $A \in \Re^{n_i \times n_i}$. Letting $A = n_s\Sigma$ produces a "generalized MEM ridge regressor" $\tilde{r} = (X^T X - n_s\Sigma + n_s\Sigma)^{-1} X^T y$. Just as ridge regression has a mean-squared-error (MSE) advantage over OLS for the classical model for noisy and/or collinear data, so this MEM ridge regressor, $\tilde{r}$, can have an MSE advantage over $\tilde{r}_{\text{MEM}}$. Define

$$\text{MSE}(v) = \mathcal{E}\left((v - r)(v - r)^T\right) \tag{6.25}$$

for $v \in \Re^{n_i}$. Note that

$$\text{MSE}(\tilde{r}_{\text{MEM}}) \rightarrow \frac{1}{n_s} M_T^{-1} (\Sigma r r^T \Sigma + \sigma_e^2 M_X) M_T^{-1} \quad \text{and} \tag{6.26}$$

$$\text{MSE}(\tilde{r}) \rightarrow \frac{1}{n_s} M_X^{-1} (\Sigma r r^T \Sigma + \sigma_e^2 M_X) X_T^{-1} + M_X^{-1} \Sigma r r^T \Sigma M_X^{-1} \tag{6.27}$$

as $n_s \rightarrow \infty$; equation 6.27 is verified in claim 6.4. A necessary and sufficient condition for $\text{MSE}(\tilde{r}) \leq \text{MSE}(\tilde{r}_{\text{MEM}})$ is developed in claim 6.5. As shown in claim 6.6, two conditions that favor $\text{MSE}(\tilde{r})$ being less than $\text{MSE}(\tilde{r}_{\text{MEM}})$ are $r^T M_T \Sigma^{-1} M_T r$ being "small" and σ_e^2 being "large." These conditions can be loosely described as "poor signal-to-noise (SNR) ratio." Notice that even if all the individual inputs have "good" SNR (the diagonal element of M_T is "large" relative the corresponding diag-

onal element of Σ), collinearity may result in $M_T\Sigma^{-1/2}$ being "small" in the crucial r direction.

While $\tilde{r}$ may be preferable to $\tilde{r}_{\text{MEM}}$ for collinear data, $\tilde{r}$ has its own well known difficulties with collinearity. Given that significance regression is a successful method for mitigating collinearity difficulties in the classical model, what is outcome of using SR for the MEM? A key difference will be in the selection of W. SR proceeds by assuming

$$\text{Var}_\infty\left(\sqrt{n_s}(\tilde{r}-p)\right) = \sigma_e^2 M_X^{-1}. \tag{6.28}$$

However as shown in claim 6.3,

$$\begin{aligned}
\text{Var}_\infty\left(\sqrt{n_s}(\tilde{r}-p)\right) &= \lim_{n_s\to\infty} n_s(\tilde{r}-p)(\tilde{r}-p)^T && (6.29)\\
&= M_X^{-1}(\Sigma r r^T \Sigma + \sigma_e^2 M_X) M_X^{-1} && (6.30)\\
&= M_X^{-1}\Sigma r r^T \Sigma M_X^{-1} + \sigma_e^2 M_X^{-1} && (6.31)
\end{aligned}$$

where p is the "optimal predictor" defined in claim 6.1. While equation 6.31 does differ from equation 6.28, the primary difference occurs in the $M_X^{-1}\Sigma r$ direction. Thus if one uses the "classical" algorithm 3.1 instead of SR-MEM, one can reasonably proceed assuming that the ignored measurement errors often will not make the SR estimate unduly unreliable: the direction selection process is only affected in one direction, and the least-squares estimate constrained to Range(W) can compare favorably to the asymptotically unbiased estimate. Further discussion of this point as it relates to scaling is given in section 8.6, while the point is demonstrated via simulation in section 9.3.

6.4 Summary

This chapter examined collinear problems where both the input and output data are corrupted by measurement noise. The general SR regression method of chapter 4 was used with the measurement error model (MEM) to produce a significance regression method appropriated for this case. This approach assumes that the input error covariances are known; often they are not. The properties and performance of regressors derived form the classical (no input noise) model were investigated for the MEM. The OLS regressor was shown to be a biased estimate of r, but an asymptotically unbiased estimate of the "optimal predictor" and a "natural ridge regressor." The SR-MEM algorithm was shown to be a modest correction to the SR method of chapter 3. These observations suggest that while the SR-MEM method is desirable when all the error covariances are known, one can usefully address collinear MEM problems using methods derived for the classical model; knowledge of the error covariances is not essential.

6.5 Appendix: Proof of Claims

This appendix presents the arguments that verify several claims made in section 6.3. Most of these claims are straightforward; they have been relegated to the appendix so that the requisite algebra would not distract from the main points. Any mention of Schneeweiß refers to [83].

Claim 6.1 *The least-squares optimal predictor is* $p = (\frac{T^T T}{n_s} + \Sigma)^{-1} \frac{T^T T}{n_s} r$.

A useful quantity to compute is

$$p = \min_{v \in \Re^{n_i}} \mathcal{E}\left((y - Xv)^T (y - Xv)\right). \tag{6.32}$$

First compute the expectation:

$$\begin{aligned} \mathcal{E}\left((y - Xv)^T (y - Xv)\right) &= \mathcal{E}\left((Tr + e - (T + E)v)^T (Tr + e - (T + E)v)\right) && (6.33) \\ &= r^T T^T T r + n_s \sigma_e^2 - 2 r^T T^T T v + v^T T^T T v && (6.34) \\ &\qquad + n_s v^T \Sigma v. \end{aligned}$$

Computing the gradient and equating it to zero,

$$\nabla_v \mathcal{E}\left((y - Xv)^T (y - Xv)\right) = -2T^T T r + 2T^T T v + 2n_s \Sigma v = 0, \tag{6.35}$$

implies

$$p = (\frac{T^T T}{n_s} + \Sigma)^{-1} \frac{T^T T}{n_s} r. \tag{6.36}$$

□

Claim 6.2 *The OLS regressor asymptotically equals the least-squares optimal predictor. That is* $lim_{n_s \to \infty} \tilde{r} = lim_{n_s \to \infty} p$.

$$\lim_{n_s\to\infty}\tilde{p} = \lim_{n_s\to\infty}(X^TX)^{-1}X^Ty \tag{6.37}$$

$$= \lim_{n_s\to\infty}\left(\frac{1}{n_s}(T^TT+T^TE+E^TT+E^TE)\right)^{-1} \left(\frac{1}{n_s}(T+E)^T(Tr+v)\right) \tag{6.38}$$

$$= M_X^{-1}M_Tr \tag{6.39}$$

$$= \lim_{n_s\to\infty} p. \tag{6.40}$$

□

Claim 6.3 $\left(\sqrt{n_s}(\tilde{r}-p)\right)$ *has an asymptotically normal distribution with zero mean and variance* $M_X^{-1}(\Sigma rr^T\Sigma+\sigma_e^2M_X)M_X^{-1}$.

This claim is an algebraic variation of a result due to Scheenweiß using the algebraic relation $y = Xr - Er + e$. Also define $E_e = [e|\ E]$ and $g = [1\ |\ -r]^T$.

$$\tilde{r}-p = (X^TX)^{-1}X^Ty - p \tag{6.41}$$

$$= \left(\frac{X^TX}{n_s}\right)^{-1}\left(\frac{X^TX}{n_s}r - \frac{X^TE}{n_s}r + \frac{X^Te}{n_s} + \Sigma r - \Sigma r\right) - p \tag{6.42}$$

$$= \left(\frac{X^TX}{n_s}\right)^{-1}\left(\frac{T^TT}{n_s}r + \frac{E^TT}{n_s}r + \frac{T^TE}{n_s}r + \frac{E^TE}{n_s}r - \frac{T^TE}{n_s}r - \frac{E^TE}{n_s}r + \frac{T^Te}{n_s} + \frac{E^Te}{n_s} + \Sigma r - \Sigma r\right) - p \tag{6.43}$$

$$= \left(\frac{X^TX}{n_s}\right)^{-1}\left(\frac{T^TE_e}{n_s}g + \frac{E^TE_e}{n_s}g + \Sigma r\right) + \left(\frac{X^TX}{n_s}\right)^{-1}\left(\frac{T^TT}{n_s} + \frac{E^TT}{n_s} + \frac{T^TE}{n_s} + \frac{E^TE}{n_s} - \Sigma\right)r - p. \tag{6.44}$$

Now

$$\lim_{n_s\to\infty}\left(\frac{X^TX}{n_s}\right)^{-1}\left(\frac{T^TT}{n_s} + \frac{E^TT}{n_s} + \frac{T^TE}{n_s} + \frac{E^TE}{n_s} - \Sigma\right)r \tag{6.45}$$

$$= M_X^{-1}(M_T+\Sigma)r - M_X^{-1}\Sigma r$$

$$= r - M_X^{-1}\Sigma r \tag{6.46}$$

$$= \lim_{n_s \to \infty} \tilde{r} \tag{6.47}$$

$$= \lim_{n_s \to \infty} p. \tag{6.48}$$

When equation 6.44 is pre-multiplied by $\sqrt{n_s}$, the second line of equation 6.44 still asymptotically vanishes due to (A8). Moreover, the $\left(\frac{T^T E_e}{n_s} g + \frac{E^T E_e}{n_s} g + \Sigma r\right)$ portion of equation 6.44 is identical to the $(I_n \otimes \gamma')\text{col}\left[\frac{1}{T}\mathcal{W}'\Xi + \frac{1}{T}\mathcal{W}'V - \Sigma_{\mathcal{W}V}\right]$ term of equation 4.1 of Schneeweiß. From here one follows Schneeweiß's development to verify the claim. □

Claim 6.4 *As* $n_s \to \infty$, $\text{MSE}(\tilde{r}) \to \frac{1}{n_s} M_X^{-1}(\Sigma r r^T \Sigma + \sigma_e^2 M_X) M_X^{-1} + M_X^{-1} \Sigma r r^T \Sigma M_X^{-1}$.

This claim is a direct extension of claim 6.3.

$$\text{MSE}(\tilde{r}) = \mathcal{E}\left((\tilde{r} - r)(\tilde{r} - r)^T\right) \tag{6.49}$$

$$= \mathcal{E}\left((\tilde{r} - r - M_X^{-1}\Sigma r + M_X^{-1}\Sigma r)(\tilde{r} - r - M_X^{-1}\Sigma r + M_X^{-1}\Sigma r)^T\right) \tag{6.50}$$

$$= \frac{1}{n_s}\mathcal{E}\left(n_s(\tilde{r} - r + M_X^{-1} r)(\tilde{r} - r + M_X^{-1} r)^T\right) - M_X^{-1}\frac{1}{\sqrt{n_s}}\mathcal{E}\left(\sqrt{n_s}(\tilde{r} - r - M_X^{-1} r)^T\right)$$
$$-\frac{1}{\sqrt{n_s}}\mathcal{E}\left(\sqrt{n_s}(\tilde{r} - r + M_X^{-1}\Sigma r)\right) r^T \Sigma {M_X^{-1}}^T + M_X^{-1}\Sigma r r^T \Sigma {M_X^{-1}}^T. \tag{6.51}$$

As $n_s \to \infty$,

$$\mathcal{E}\left((\tilde{r} - r + M_X^{-1}\Sigma r)(\tilde{r} - r + M_X^{-1}\Sigma r)^T\right) \to \frac{1}{n_s} M_X^{-1}(\Sigma r r^T \Sigma + \sigma_e^2 M_X) M_X^{-1} \tag{6.52}$$

and

$$\mathcal{E}\left(\sqrt{n_s}(\tilde{r} - r + M_X^{-1} r)^T\right) \to 0 \tag{6.53}$$

by the same argument used in claim 6.3, so the claim holds. □

Claim 6.5 *For large but finite* n_s $\text{MSE}(\tilde{r}_{\text{MEM}}) - \text{MSE}(\tilde{r})$ *is a positive semi-definite matrix.*

This claim does not hold for all problems, but does hold for many collinear and/or noisy problems. The following "proof" develops a (tautological) necessary and sufficient condition for the claim to hold.

Defining $X_D = XD_1$ for any non-singular D_1,

$$\tilde{r}(X_D, y) = D_1^{-1}\tilde{r}(X, y) \text{ and} \tag{6.54}$$

$$\tilde{r}_{\text{MEM}}(X_D, y) = D_1^{-1}\tilde{r}_{\text{MEM}}(X, y). \tag{6.55}$$

Thus both $\tilde{r}$ and $\tilde{r}_{\text{MEM}}$ are scale invariant. Recall that for any two matrices A and B of the same dimensions

$$A \geq B \quad \Leftrightarrow \quad DAD \geq DBD \tag{6.56}$$

for any non-singular symmetric D where "$A \geq B$" means $A - B$ is a positive semi-definite matrix. Choose $D_1 = \Sigma^{-\frac{1}{2}}$. Then the rescaled problem has $\Sigma_D = I$. Under this scaling, as $n_s \to \infty$

$$\text{MSE}(\tilde{r}_{\text{MEM}}) \to \frac{1}{n_s} M_T^{-1}(rr^T + \sigma_e^2 M_X^{-1})M_T^{-1}, \text{ and} \tag{6.57}$$

$$\text{MSE}(\tilde{r}) \to \frac{1}{n_s} M_X^{-1}(rr^T + \sigma_e^2 M_X)M_X^{-1} + M_X^{-1}rr^T M_X^{-1}. \tag{6.58}$$

Thus for n_s large,

$$\begin{aligned} \text{MSE}(\tilde{r}_{\text{MEM}}) - \text{MSE}(\tilde{r}) &= \frac{1}{n_s}\Big(M_T^{-1}(rr^T + \sigma_e^2 M_X^{-1})M_T^{-1} - \\ &\qquad M_X^{-1}(rr^T + errM_X^{-1})M_X^{-1}\Big) - M_X^{-1}rr^T M_X^{-1} \qquad (6.59) \\ &= \frac{1}{n_s} M_T^{-1}rr^T M_T^{-1} - (1 + \frac{1}{n_s})M_X^{-1}rr^T M_X^{-1}. \\ &\qquad + \frac{\sigma_e^2}{n_s}\left(M_T^{-1}M_X M_T^{-1} - M_X^{-1}\right). \qquad (6.60) \end{aligned}$$

The third term (with the $\frac{\sigma_e^2}{n_s}$ co-efficient) of equation 6.60 is always positive definite. However the sum of the first two terms is an indefinite matrix: these terms may sum to the null matrix, to a rank one matrix with either a positive or a negative eigenvalue, or to a rank two matrix with one positive and one negative eigenvalue. The positive definite third term may or may not overcome the negative eigenvalue, depending on the values of the parameters; therefore the right-hand side (RHS) of equation 6.60 may or may not be positive semi-definite. The claim is true if and only if the RHS of equation 6.60 is positive semi-definite for appropriately rescaled data. □

Claim 6.6 *Claim 6.5 holds for any large but given* n_s *if* $\frac{\sigma_e^2}{n_s+1} \geq \sum_{i=1}^{n_i} \frac{\rho_i^2 \lambda_i^2}{(\lambda_i+1)(2\lambda_i+1)}$, *where* v_i *are the eigenvalues of* $\Sigma^{-\frac{1}{2}} M_T \Sigma^{-\frac{1}{2}}$ *with corresponding eigenvalues* $\lambda_i > 0$ *and* $\rho_i = v_i^T \Sigma^{\frac{1}{2}} r$.

A sufficient condition for claim 6.5 to hold is

$$\frac{\sigma_e^2}{n_s} \left(M_T^{-1} M_X M_T^{-1} - M_X^{-1} \right) \geq (1 + \frac{1}{n_s}) M_X^{-1} r r^T M_X^{-1} \tag{6.61}$$

when the data have been rescaled such that $\Sigma = I$. Since the LHS of equation 6.61 is positive definite, Farebrother's 1976 result (Theorem 2.5.2 of [24]) reveals that equation 6.61 holds if and only if

$$\frac{\sigma_e^2}{n_s + 1} \geq r^T M_X^{-1} \left(M_T^{-1} M_X M_T^{-1} - M_X^{-1} \right)^{-1} M_X^{-1} r. \tag{6.62}$$

Since the v_i are also the eigenvalues of M_X with corresponding eigenvalues $\lambda_i + 1$, one can diagonalize the matrices and substitute $M_T = \Lambda$ and $M_X = \Lambda + I$ where

$\Lambda = \text{Diag}(\lambda_1, \ldots, \lambda_{n_i})$. Then

$$\frac{\sigma_e^2}{n_s+1} \geq \begin{bmatrix} \rho_1 \\ \vdots \\ \rho_{n_i} \end{bmatrix}^T (\Lambda + I)^{-1} \left(\Lambda^{-1}(\Lambda + I)\Lambda^{-1} - (\Lambda + I)^{-1}\right)^{-1} (\Lambda + I)^{-1} \begin{bmatrix} \rho_1 \\ \vdots \\ \rho_{n_i} \end{bmatrix} \tag{6.63}$$

$$\frac{\sigma_e^2}{n_s+1} \geq \begin{bmatrix} \rho_1 \\ \vdots \\ \rho_{n_i} \end{bmatrix}^T (\Lambda + I)^{-1} \left(\text{Diag}\left(\frac{2\lambda_i + 1}{(\lambda_i + 1)\lambda_i^2}\right)\right)^{-1} (\Lambda + I)^{-1} \begin{bmatrix} \rho_1 \\ \vdots \\ \rho_{n_i} \end{bmatrix} \tag{6.64}$$

$$\frac{\sigma_e^2}{n_s+1} \geq \sum_{i=1}^{n_i} \frac{\rho_i^2 \lambda_i^2}{(\lambda_i + 1)(2\lambda_i + 1)}. \tag{6.65}$$

Notice that inequality 6.65 will hold when λ_i is "small" for any direction for which ρ_i "large," or , equivalently, when $r^T M_T M_T r$ is "small" for properly scaled data ($D_1 = \Sigma^{-\frac{1}{2}}$). For unscaled data this condition is

$$\frac{\sigma_e^2}{n_s+1} \geq r^T M_T \Sigma^{-1} M_T r. \tag{6.66}$$

□

Chapter 7

Robust Significance Regression

7.1 Introduction

The development thus far has assumed that ordinary least squares is the desired objective function. However, one may also wish to address weighted least squares objective functions. Thus the weighted least squares significance regression algorithm is developed. Next the issue of robustness is addressed. Here robustness is defined to be tolerance for outliers and/or deviations from the assumed distribution. Significance regression is shown to have poor robustness properties; however robustness can be achieved by choosing a better objective function than least-squares. Such regressors are call M-estimators and possess both a strong theoretical foundation and a successful history of practical use. The M-regressors can be expressed using a weighted least-squares objective function but are unreliable for collinear data. In this chapter a robust significance regressor is developed using M-estimation to generate the weights for the weighted least-squares significance regression method. The resulting restriction regressor inherits the robustness properties of the M-estimator while maintaining SR's ability to treat collinearity.

7.2 Weighted Least Squares Significance Regression

Consider again the classical linear model

$$y = Xr + e \tag{7.1}$$

but where $\mathcal{E}\left(ee^T\right) = \sigma_e^2 P$ for some non-singular symmetric P. Moreover, consider the weighted least-squares problem,

$$\begin{aligned} \tilde{r}_{\text{WLS}} &= \arg\min_{v \in \Re^{n_i}} (y - Xv)^T M (y - Xv) && (7.2) \\ &= (X^T M X)^{-1} X^T M y && (7.3) \end{aligned}$$

for some non-singular M. Additionally assume that M is symmetric positive-definite; most physically interesting weighting matrices will conform to these additional assumptions. As is well-known, $\tilde{r}_{\text{WLS}}$ is the minimum-variance unbiased estimator of r when $P = M^{-2}$. However, one may not always be interested in the minimum-variance unbiased estimator (as in the robust regressors developed below), so this problem is worth studying in its own right. Notice that $\tilde{r}_{\text{WLS}}$ maintains the same weakness for collinearity as $\tilde{r}$.

Following the development in chapter 4, the null hypothesis of interest is $< w, r > = 0$ for which a natural test statistic in the weighed least squares context is

$$\tau(w, y) = \frac{(w^T \tilde{r}_{\text{WLS}})}{\sqrt{\text{Var}(w^T \tilde{r}_{\text{WLS}})}}. \tag{7.4}$$

Computing the necessary variance

$$\text{Var}(w^T \tilde{r}_{\text{WLS}}) = w^T \mathcal{E}\left((r_{\text{WLS}} - r)(r_{\text{WLS}} - r)^T\right) w \tag{7.5}$$

$$= w^T \mathcal{E}\left((X^T M X)^{-1} X^T M e e^T X (X^T M X)^{-1}\right) w \tag{7.6}$$

$$= \sigma_e^2 w^T (X^T M X)^{-1} X^T M P X (X^T M X)^{-1} w \tag{7.7}$$

leads to

$$\tau(w, y) = \frac{w^T \tilde{r}_{\text{WLS}}}{\sqrt{w^T (X^T M X)^{-1} X^T M P X (X^T M X)^{-1} w \sigma_e^2}}. \tag{7.8}$$

From this test statistic the weighted least squares significance regression algorithm is:

Algorithm 7.1 (SR-WLS)

$$\tilde{r}_{\text{WLS}} = \quad (X^T M X)^{-1} X^T M y \tag{7.9}$$

$$\text{Var}(\tilde{r}_{\text{WLS}}) = \quad (X^T M X)^{-1} X^T M P M X (X^T M X)^{-1} \tag{7.10}$$

$$W_0 = \quad 0, \quad W_0 \in \Re^{n_s} \tag{7.11}$$

$$\text{DO} \quad i = \quad 1, n_d$$

$$v = \quad (I - W_{i-1} W_{i-1}^T)\left(\text{Var}(\tilde{r}_{\text{WLS}})\right)^{-i} \tilde{r}_{\text{WLS}} \tag{7.12}$$

$$w_i^{opt} = \quad \frac{v}{||v||} \tag{7.13}$$

$$W_i = \quad [w_1^{opt} | w_2^{opt} | \cdots | w_i^{opt}] \tag{7.14}$$

END DO

$$\tilde{b}_{\text{WLS}} = \quad W_{n_d} (W_{n_d}^T X^T M X W_{n_d})^{-1} W_{n_d}^T X^T M y. \tag{7.15}$$

The SR-WLS regressor, $\tilde{b}_{\text{WLS}}$, is similar to performing SR on data that has been scaled by $D_2 = M^{\frac{1}{2}}$ (see section 8.2). That is, let $X_M = M^{\frac{1}{2}} X$ and $y_M = M^{\frac{1}{2}} y$, and then perform normal SR using X_M and y_M. The distinction appears in equations 7.10 and 7.12. Under such a scheme, the variance described in equation 7.10 would be the same, but would not be used in equation 7.12. Instead, equation 7.12 would be replaced by

$$w_i^{opt} = (I - W_{i-1} W_{i-1}^T)(X^T M X)^i \tilde{r}_{\text{WLS}}. \tag{7.16}$$

Thus the "rescaled SR" approach loses all information about P; if the rescaled errors $(M^{\frac{1}{2}}e)$ are not homoscedastic and independent, the method will suffer.

Once again, the issue of how to choose n_d arises. One can use cross-validation [87] and the WLS objective in equation 7.2. Alternatively, one could pursue a statistical test. $\tau(w,y)$ is normally distributed if e is normal, $\mathcal{H}_0^1$ of chapter 3 holds, and w is independent of y. Unfortunately $(\tau_i^{opt}(y))^2$ has a more involved distribution. Also, the motivation for using a weighting matrix $M \neq P^{-1}$ will probably also impact how one choose n_d. For example, one typically tries to use the χ^2 distribution when building a statistical test for choosing n_d. However, a major motivation for developing the SR-WLS algorithm is to lay the foundation for robust significance regression — and the χ^2 distribution is a poor description of data with outliers. Thus, the issue of choosing n_d is left open for moment.

SR-WLS has applications beyond being a step in the derivation of a robust SR method. The least-squares objective works with an absolute error description; however in many applications one is interested in minimizing a relative error objective. If one is building a predictor of chemical composition, a predictor that provides excellent precision of the concentration of the most common chemical at the expense of predictive ability for the trace chemicals will not always reflect the practitioner's desires. If the y values were themselves corrupted by relative errors, then the data would be heteroscedastic and the scaling suggestion to be discussed in section 8.5 would apply. However, if the errors really were homoscedastic, one could employ a weighted least-squares objective using weights computed from y. Letting $y = [\psi_1, \psi_2, \ldots \psi_{n_s}]^T$, the least-relative-squared error objective is

$$\min_{v \in \Re^{n_i}} \sum_{j=1}^{n_s} \left(\frac{\psi_j - x_i^T v}{\psi_j} \right)^2 \tag{7.17}$$

which immediately gives rise to the diagonal weighting matrix M with the jth diag-

onal component of $1/\psi_j$.

7.3 Robust Regression

Robust estimation is supported by a rich and successful corpus of theory; only brief portions of the theory needed to develop a robust significance regressor are touched upon here. The interested reader is referred to Huber [40, 42] and Tukey [33] for further development, as the claims made here are derived from these sources. Some might object that these approaches are not particularly relevant: since most practitioners will examine their data and reject obvious outliers, why develop outlier resistant regressors? Huber answers this question by claiming "only the best among these rejection rules can more or less compete with other good robust estimators... Rules based on the studentized range, for example, are disastrously bad [42, p. 3]."

To address robustness, two questions need to be addressed: What is "robustness," and how does one measure it? In this development a limited definition of robustness is used: a regressor is robust if it (1) is insensitive to small deviations of the error distribution from the assumed distribution and (2) remains bounded in the face of small numbers (less than 30%, say) unbounded gross errors. Quantifying the first part of the robustness definition is beyond the scope of the current effort; chapter 2 of Huber [42] is good starting point. However the point can be briefly illustrated using Tukey's famous example [69]. Consider the problem of estimating the variance of a distribution. Consider a sef of data $\{\psi_1, \ldots, \psi_{n_s}\}$. Two common measures of scatter are the mean absolute deviation

$$\sigma_1 = \frac{1}{n} \sum_{i=1}^{n_s} |\psi_i - \bar{\psi}| \tag{7.18}$$

and the mean square deviation

$$\sigma_2 = \sqrt{\frac{1}{n}\sum_{i=1}^{n_s}(\psi_i - \bar{\psi})^2} \tag{7.19}$$

where $\bar{\psi}$ is the sample mean. For normally distributed errors, σ_2 is the "efficient" estimator in the sense of the Cramer-Rao bound. However, consider the case where all samples are independently drawn from normal distributions with the same mean, but with probability ϵ an observation is drawn from a "bad" normal distribution with three times the deviation (square root of the variance). To compare the two estimators, use the asymptotic relative efficiency,

$$\mathrm{ARE}(\epsilon) = \lim_{n\to\infty} \frac{\mathrm{Var}(\sigma_2)/\left(\mathcal{E}\left(\sigma_2\right)\right)^2}{\mathrm{Var}(\sigma_1)/\left(\mathcal{E}\left(\sigma_1\right)\right)^2}. \tag{7.20}$$

Of course, for $\epsilon = 0$ the ARE is less than one: $\mathrm{ARE}(0) = 0.876$. This is the source of the common claim that the mean square deviation is "12% more efficient." However for $0.002 < \epsilon < 0.998$ $\mathrm{ARE}(\epsilon) > 1$; if only two observations per thousand deviate from the normality assumption underlying the σ_2 claim to efficiency, σ_2 is *not* the efficient estimator. Thus, σ_2 is seen to have poor robustness in the first sense. In fact, many maximum likelihood estimators derived assuming a normal distribution for the errors have poor robustness in this sense.

The second portion of the robustness definition is quantified via the concept of the "breakdown point." As defined by Hampel [26], the breakdown point of an estimator is the largest possible fraction of the observations for which there is a bound on the change in the estimate when that fraction of the sample is altered without restriction. Thus, a necessary condition for robustness is a non-zero breakdown point. Notably, most of the estimators derived from the common (unweighted) least squared error objective, including $\tilde{r}$ of equation 2.4 and $\tilde{b}$ of equation 2.8 have a zero breakdown point: if any given observation is altered without bound, the estimator is also altered

without bound.

There are numerous robust estimators to choose from; here the class examined is the M-estimators. M-estimators minimize the maximal asymptotic variance over the relevant family of distributions when such a statement is meaningful [41], typically have high breakdown points [33], and have been shown to have superior robust regression performance. [42]. Quoting from Goodall [22],

> The M-estimate $\mu_{n_s}(\psi_1, \ldots, \psi_{n_s})$ for the function $\rho(\cdot\,; \mu)$ and the sample $\psi_1, \ldots, \psi_{n_s}$ is the value of μ that minimizes the objective function $\sum_{j=1}^{n_s} \rho(\psi_j; \mu)$.

A necessary condition M-estimators must satisfy is

$$\frac{\partial \sum_{j=1}^{n_s} \rho(\psi_j; \mu)}{\partial \mu} = 0, \tag{7.21}$$

which can be expressed as

$$\sum_{j=1}^{n_s} \Psi(\psi_j; \mu) = 0 \tag{7.22}$$

where $\Psi(\psi_j; \mu) = \frac{\partial \rho(\psi; \mu)}{\partial \mu}\Big|_{\psi=\psi_j}$; typically $\rho(\psi_j; \mu)$ is differentiable with respect to μ almost everywhere.

The two most familiar M-estimates are the sample mean, derived from $\rho_1(\psi_j; \mu) = \frac{1}{2}(\psi_j - \mu)^2$, and the sample median, derived from $\rho_2(\psi_j; \mu) = |\psi_j - \mu|$. Importantly, any M-estimator can be redescribed by the weighted least squares objective $\rho_3(\psi_j; \mu) = \frac{\omega_j}{2}(\psi_j - \mu)^2$ where $\omega_j = \frac{\Psi(\psi_j; \mu)}{\mu - \psi_j}$. To develop robust regressors, an additional constraint will be added to the functional form of ρ: $\rho(\psi_j; \mu) = \rho(\Delta_j)$ where $\Delta_j = \psi_j - \mu$. Also, for regression, one must be able to define what is "large" and what is "small" for the sake of identifying outliers. This is done via the scale parameter σ_{robust}, which itself is a robust variance estimator. Defining the j-th observation of the output (y) to be

ψ_j, the unique robust regressor $\tilde{r}_{\text{robust}}$ and scale parameter σ_{robust} is defined by

$$\sum_{j=1}^{n_s} \Psi\left(\frac{\Delta_j}{\sigma_{\text{robust}}}\right) = 0 \quad \text{and} \quad \sum_{j=1}^{n_s} \xi\left(\frac{\Delta_j}{\sigma_{\text{robust}}}\right) = 0 \tag{7.23}$$

where $\xi\left(\frac{\Delta_j}{\sigma_{\text{robust}}}\right) = \frac{\Delta_j}{\sigma_{\text{robust}}}\Psi\left(\frac{\Delta_j}{\sigma_{\text{robust}}}\right) - \rho\left(\frac{\Delta_j}{\sigma_{\text{robust}}}\right)$ and $\Delta_j = \psi_j - x_j^T\tilde{r}_{\text{robust}}$.

Solving the equations 7.23 can be done via the following iteratively re-weighted least squares (IRLS) algorithm:

Algorithm 7.2 (Robust M-Regression)

$$i = 0 \tag{7.24}$$

$$\tilde{r}_0 = (X^TX)^{-1}X^Ty \tag{7.25}$$

$$\sigma_0 = \frac{\|X\tilde{r}_0 - y\|}{\sqrt{n_s - n_i}} \tag{7.26}$$

$$\Delta_j = \psi_j - x_j^T\tilde{r}_0 \quad \forall j \tag{7.27}$$

$$M_0^{\frac{1}{2}} = \text{Diag}\left(\frac{\Psi(\Delta_j/\sigma_0)}{\Delta_j/\sigma_0}\right) \tag{7.28}$$

DO

$$i = i + 1 \tag{7.29}$$

$$\tilde{r}_i = (X^TM_{i-1}X)^{-1}X^TM_{i-1}y \tag{7.30}$$

$$\Delta_j = \psi_j - x_j^T\tilde{r}_i \quad \forall j \tag{7.31}$$

$$\sigma_i = \frac{\sigma_{i-1}}{(n_s - n_i)\alpha}\sqrt{\sum_{j=1}^{n_s} \xi\left(\frac{\Delta_j}{\sigma_{i-1}}\right)} \tag{7.32}$$

$$M_i^{\frac{1}{2}} = \text{Diag}\left(\frac{\Psi(\Delta_j/\sigma_i)}{\Delta_j/\sigma_i}\right) \tag{7.33}$$

UNTIL convergence

$$M_{\text{robust}} = M_i \tag{7.34}$$

$$\sigma_{\text{robust}} = \sigma_i \tag{7.35}$$

$$\tilde{r}_{\text{robust}} = r_i \tag{7.36}$$

where α is the $\mathcal{E}(\xi(\cdot))$ when the argument is a unit normal random variable. For Huber's commonly-used "proposition 2" objective function (described below) $\alpha = 0.258$. This algorithm is adapted from Huber's `Algorithm H` [42] and has been proven to converge uniquely. In fact, this algorithm converges quickly, typically in less than ten iterations.

An open question is the choice of $\rho(\cdot)$. A common and successful [11] M-estimator uses Huber's "proposition 2" objective function:

$$\rho\left(\frac{\Delta_j}{\sigma_{\text{robust}}}\right) = \begin{cases} \frac{\Delta_j^2}{2\sigma_{\text{robust}}^2} & \text{for } |\frac{\Delta_j}{\sigma_{\text{robust}}}| \leq 1 \\ |\frac{\Delta_j}{\sigma_{\text{robust}}}| - \frac{1}{2} & \text{for } |\frac{\Delta_j}{\sigma_{\text{robust}}}| > 1 \end{cases} . \tag{7.37}$$

One can see that the resulting M-estimator uses the mean, which is "efficient" (in the classical sense) but not robust, when the error is "small," and uses the median, which is robust but not "efficient," when the error is "large." In fact, the "proposition 2" estimator is the maximum likelihood estimate when the samples are independently drawn from a normal distribution contaminated by the "unique symmetric asymptotically least favorable distribution [40]." For particularly heavy-tailed error distributions, a better $\rho(\cdot)$ is Tukey's biweight [33]:

$$\rho\left(\frac{\Delta_j}{\sigma_{\text{robust}}}\right) = \begin{cases} \frac{1}{6}\left(1 - \left(1 - \frac{\Delta_j^2}{\sigma^2}\right)^3\right) & \text{for } |\frac{\Delta_j}{\sigma_{\text{robust}}}| \leq 1 \\ \frac{1}{6} & \text{for } |\frac{\Delta_j}{\sigma_{\text{robust}}}| > 1 \end{cases} . \tag{7.38}$$

One should keep in mind that the theory supporting M-regressors assumes independent and *a priori* homoscedastic errors; therefore one should always scale that data such that $P = \sigma_e^2 I$ before using algorithm 7.2. More importantly, M-estimators in general, and these two in particular, are *not* the minimum-variance unbiased estimators; achieving the minimum-variance property comes at the direct expense of robustness. However, both the "proposition 2" and Tukey's biweight objective func-

tions lead to unbiased estimators with breakdown points of almost 0.5 [22].

7.4 Robust Significance Regression

While the algorithm 7.2 produces robust regressors, it inherits OLS's weakness for collinear problems. As shown in the variance,

$$\mathrm{Var}(\tilde{r}_{\mathrm{robust}}) = (X^T M_{\mathrm{robust}} X)^{-1} X^T M_{\mathrm{robust}} P M_{\mathrm{robust}} X (X^T M_{\mathrm{robust}} X)^{-1}, \qquad (7.39)$$

the estimate will often be unreliable in the directions for which the eigenvalues of $X^T M X$ are small. On the other hand, SR, which does treat collinear problems successively, has a breakdown point of zero and is therefore not robust. The next step is to combine the two desirable properties; the clear and simple manner to do this is via SR-WLS.

Algorithm 7.3 (SR - Robust)

1. Rescale data such that $P = \sigma_e^2 I$.
2. Choose an M-estimation objective function, $\rho(\cdot)$,
3. Compute M_{robust} and σ_{robust} with algorithm 7.2.
4. Use algorithm 7.1 to generate $\tilde{b}_{\mathrm{robust}}$.

Notice that the resulting restriction regressor,

$$\tilde{b}_{\mathrm{robust}} = \arg \min_{v \in \mathrm{Range}(W_{n_d})} (Xv - y)^T M (Xv - y) \qquad (7.40)$$

has the same breakdown point as $\tilde{r}_{\text{robust}}$ and inherits *all* of the robustness properties of the M-estimator when $r \in \text{Range}(W_{n_d})$. Thus using significance regression does not cause any "loss of robustness" but does maintain the ability to treat collinearities.

Algorithm 7.1 does not firmly specify how to choose n_d. As mentioned in section 7.1, one could use cross-validation with the robust objective function in equation 7.40. However, one can also develop a useful alternative based on the significance regression framework. Following the approach in Chapter 3 is useful in broad stroke, but dubious in detail; the "statistical" method put forward in section 3.2 for choosing n_d relies on the χ^2 distribution, which is known to be a poor description of data that contains outliers and or deviates from the normal distribution. As noted in section 3.2, and demonstrated in chapter 9, a very simple and effective decision rule is: if $(\tau(w_i^{opt}, y))^2 > n_p)$ then $n_d \geq i$. Use of this decision rule directly can still be dangerous, since one typically employs $\tilde{\sigma}_e^2$, which is not robust. However, using the scale parameter σ_{robust} allows one to define a "robust" significance regression objective function

$$(\tau_{\text{robust}}(w, y))^2 = \frac{(w^T \tilde{r}_{\text{robust}})^2}{w^T (X^T X)^{-1} w \sigma^2_{\text{robust}}} \tag{7.41}$$

from which one can use the decision rule: if $(\tau_{\text{robust}}(w_i^{opt}, y))^2 > n_p)$ then $n_d \geq i$. This decision rule does not rest on theoretical derivation; in particular, it is not equivalent to the 50% F-test for large n_s. However, as will be demonstrated in chapter 9, it does provide a useful *ad hoc* rule.

7.5 Summary

This chapter developed a novel robust restriction regressor, SR-Robust. This regressor employs the objective functions that make the M-estimators tolerant of outliers and distributionally robust while using the SR method to treat collinearities. By choosing among the well-analyzed M-estimation objective functions, one can "tune"

the method if one knows that the error distribution is "heavy-tailed" or if one has other special knowledge of the error distribution. The computation of the SR-Robust regressor involves using the weighted-least-squares significance regression algorithm (SR-WLS), also developed in this chapter. SR-WLS can be used to address other objectives, such as buidling predictors that minimize relative error.

Chapter 8

Scaling for Significance Regression

8.1 Introduction

An essential step in the use of any biased regressor is the scaling of the data; unlike the unbiased regressor $\tilde{r}$, biased regressors are profoundly affected by scaling. Despite this importance, scaling methods for restrictions regressors are inadequately understood. The first section illustrates the importance of scaling and states some common objectives for scaling methods. Next, the close link between scaling and one's *a priori* beliefs is examined. The final three sections develop scaling methodologies to address the stated objectives.

8.2 Scaling Objectives

All "real" data is always scaled: whether the quantity in question is a length, a time interval, or a percentage, measured quantities always carry some type of unit. The choice of units can have a dramatic impact on biased regressors. Consider, for example, significance regression. Let z_i be the vector of inputs for the i-th input; that is, let $X = [z_1|\dots|z_{n_i}]$. Then $w_1^{opt} = X^T y = [z_1^T y \ z_2^T y \dots z_{n_i}^T y]^T$. The "most

significant" vector is formed from the covariances between the individual inputs and the output. Herein one can see the effect of scaling; if one of the z_i is multiplied by a large constant, say by choosing units of microns instead of kilometers, then that input will figure much more prominently in the "most significant" vector. The problem is pervasive; most other biased regressors, including ridge regression [85] and PCR [45], are just as sensitive. Clearly, one would like to mollify such effects. The "mollification" is typically done by scaling the data, as in

$$X = D_2 X_{raw} D_1 \tag{8.1}$$

$$y = D_2 y_{raw} \tag{8.2}$$

where $D_1 \in \Re^{n_i \times n_i}$ and $D_2 \in \Re^{n_s \times n_s}$ are non-singular and referred to as the input scaling and output scaling, respectively. In this study, D_1 and D_2 are also assumed to be symmetric and positive definite; this assumption can be relaxed, but is met in virtually all scalings used in practice.

The primary objectives of most scaling regimes are

1. to be computationally inexpensive,
2. to remove "scaling" (measurement unit) effects from the raw data,
3. to account for correlated and heteroscedastic output errors (different samples having different error variances), and
4. to mitigate the effects of input measurement errors.

The first objective is a matter of practicality; if computing a "good" scaling requires an order of magnitude more computations than the regression itself, then one must question if the scaling is an aid for the regression process or a replacement. For this reason scaling matrices are usually diagonal matrices computed from readily available quantities. The second and third objectives can be adequately addressed with existing

theory. The fourth objective is almost always approached via heuristics; scant theory exists to guide the user. After discussing several general aspects of scaling, each of these objectives will be addressed in turn below.

8.3 Aspects of Scaling

The meaning of scaling for ridge regressors is well understood: the scaling incorporates one's *a priori* beliefs about r. Using an explicitly Bayesian interpretation, if one believes *a priori* that r has a distribution with mean 0 and dispersion A^{-1}, then the Bayesian (or ridge) regressor is [24]

$$\tilde{b} = \arg\min_{b \in \Re^{n_i}} \|y - Xb\| + b^T A b. \tag{8.3}$$

If scaling D_1 is applied to both the data and the prior, then

$$\begin{aligned} \tilde{b} &= D_1 \arg\min_{b \in \Re^{n_i}} \|y - XD_1 b\| + b^T D_1 A D_1 b & (8.4) \\ &= \arg\min_{b \in \Re^{n_i}} \|y - Xb\| + b^T A b; & (8.5) \end{aligned}$$

the Bayesian regressor is scale invariant. However, if one scales the data only, then

$$\begin{aligned} \tilde{b} &= D_1 \arg\min_{b \in \Re^{n_i}} \|y - XD_1 b\| + b^T A b & (8.6) \\ &= \arg\min_{b \in \Re^{n_i}} \|y - Xb\| + b^T D_1^{-1} A D_1^{-1} b. & (8.7) \end{aligned}$$

The scaling changes the prior; for most ridge regressors, $A = I$ and the scaling *is* the prior [61]. As shown in chapter 2, ridge regressors and restriction regressors arise from applying different but related styles of constraints to the same optimization problem; thus, the scaling for ridge restriction regressors also reflects the user's *a priori* beliefs about r.

One should be wary of devoting too much effort to computing "optimal" scalings; as shown through the Bayesian description, the choice of scaling and the regression are inseparable. For example, a scaling that always meets the second objective and that is provably "optimal" for certain squared-error objective functions is $D_1 = (X^T X)^{-1}$. Under this scaling, $w_1^{opt} = D_1 X^T y = (X^T X)^{-1} X^T y = \tilde{r}$. By use of a particular scaling, significance regression has been contorted to exactly mimic OLS. This will rarely be regarded as "improvement." Also, given the considerable computational cost of computing the matrix inverse, this scaling fails the first objective. The remainder of this study will reflect current scaling practice and focus on diagonal D_1.

Before moving on to the main results, one peculiar form of scale invariance for SR is noted. The above scaling description used the regression viewpoint. From the factor analysis viewpoint, w_1^{opt} is interpreted as a basis vector for the x_i. Thus, to "unscale" the first significant vector is to perform a change of basis back to the original basis:

$$w_1^{opt} = D_1^{-1}(X_{raw} D_1)^T y_{raw} = X_{raw}^T y_{raw}. \tag{8.8}$$

Thus the first significant vector for factor analysis is scale invariant. However, the scaling chosen does impact the orthogonality constraint, so the later significant directions are not scale invariant.

8.4 Removing the Effect of Units

As noted in the introduction, a major motivation for scaling is to remove the effect of different units for the various inputs. This objective can be concisely stated as follows: one desires an input scaling regime, say the matrix function $D(\cdot)$, such that for any diagonal positive A

$$X = X_{raw}\ D(X_{raw} A) \tag{8.9}$$

yields the same X. This objective proves remarkably easy to meet. The two most popular methods are auto-scaling (AS) and significance scaling (SS). AS involves dividing each z_i by the standard deviation of that z_i; that is, D_1^{-2} equals the diagonal elements of X^TX. With AS the "most significant" vector is formed from the correlation co-efficients between the individual inputs and the output, providing a heuristic motivation for using autoscaling when the inputs are uncorrupted by measurement noise. Indeed, many leading authors use AS ("standardized variables") without comment or justification [58, 88].

When the inputs are known to be corrupted by noise, another common scaling is SS [62]. When the input measurement errors are independent and homoscedastic, SS involves dividing the z_i by the standard deviation of measurement error for that input. Since a change of units multiplies both the signal and the noise by the same constants, SS also meets the objective for removing the effect of units. Thus, the two most popular input scalings meet the first and second objectives. However, as will be shown in sections 8.6 and 9.5, the scaling approaches differ markedly when input measurement errors are present.

8.5 Accounting for Heteroscedasticity

Often not all of the samples will have the equal error covariances or will not be independent. For the classical model, equation 2.1, this is mathematically stated as $\mathcal{E}\left(ee^T\right) = \sigma_e^2 P$ where $P \in \Re^{n_s \times n_s}$ is symmetric and non-singular. For the classical case, the minimum variance unbiased estimator, $\tilde{r}$ has long been know to result from

$$D_2 = P^{-\frac{1}{2}}. \tag{8.10}$$

Since the SR algorithm explicitly uses $\tilde{r}$, one should use this scaling whenever the assumptions of the classical model hold.

The other form of heteroscedasticity is more difficult. All of the results in chapter 6 for the measurement error model were derived under the assumption of homoscedasticity; however, the input errors may be heteroscedastic. For example, an experimenter may improve her/his techniques with experience, so later measurements will be more precise. If one faces this problem, analysis of SR-MEM and SR-Robust suggests two possible approaches; neither is rigorous. The first approach is to rely on robust regression. As briefly discussed in chapter 7 and elaborated in Huber [42] and Hoaglin *et al.*[33], robust estimators are robust to both outliers and perturbations of the distribution. If one views the heteroscedastic samples as being drawn from the perturbing distribution, then use of algorithm 7.3 is reasonable.

The second approach insists on using the results for the classical model. Recall the measurement error model from chapter 6:

$$y = Tr + e, \quad \text{and} \tag{8.11}$$

$$X = T + E. \tag{8.12}$$

These equations are algebraically equivalent to

$$y = (X - E)r + e = Xr + (e - Er). \tag{8.13}$$

Thus the measuremt error model can be made to appear similar to the classical model that drives the SR algorithm; continuing that similarity suggests that

$$D_2 = \left(\mathcal{E}\left((e - Er)(e - Er)^T\right)\right)^{-\frac{1}{2}} \tag{8.14}$$

is a useful scaling. This scaling reduces to the optimal scaling when $E = 0$, produces a scaling proportional to the identity matrix when the samples are independent and homoscedastic, and embodies the intuitive idea that "cleaner data should be given

greater weight." Of course r is unknown and must be estimated. Since the estimate of r will only affect D_2, the estimate used for r is not crucial and $\tilde{r}$ (which is input-scale invariant) is a reasonable starting point.

8.6 Mitigating Input Measurement Errors

Rigorously treating input measurement errors was detailed in chapter 6. However, common practice dictates the use of input scaling and the classical model; moreover, section 6.3 showed that the SR method for the classical model can be effective for the MEM. Thus this section will investigate input scaling as an approach to the MEM for the sake of both custom and extending the results of section 6.3. This section will use the assumptions of chapter 6, including the assumption that the measurement errors are homoscedastic and independent between samples, although the input measurement errors may be correlated for any given sample. For these assumptions, significance scaling becomes $D_1 = \Sigma^{-\frac{1}{2}}$.

As noted above, the input scaling is an embodiment of ones *a priori* beliefs about r. Thus AS can be interpreted as the belief "all inputs are equally useful," while SS can be interpreted as the belief "the usefulness of an input is inversely proportional to its measurement noise level." One should choose between AS and SS based upon which belief seems more appropriate. While one might hope for stronger theoretical support than the vague semantics proffered thus far, the literature offers little support. When Wegscheider attempted to systematically study certain scaling effects, he found "there was no accessible prior knowledge..., nor were literature recommendations available [43]." In Mejdell's successful use of a restriction regressor for control of a distillation column, he was unable to locate a rigorous motivation for any scaling scheme and ultimately designed his own scaling regime based solely on intuition derived from control theoretic experiences[63].

Given the importance of the issue and the lack of compelling theory, SR-MEM will be used to provide two theoretical "anecdotes" in support of significance scaling with significance regression whenever other *a priori* knowledge is lacking. These results are indicative rather than compelling, but do provide more theoretical guidance than is currently available. As derived in chapter 6, for n_s large

$$\mathrm{Var}(\tilde{r}) \rightarrow \sigma_e^2 (X^T X)^{-1} + n_s (X^T X)^{-1} \Sigma r r^T \Sigma (X^T X)^{-1}. \tag{8.15}$$

However in SR one only makes use of the full rank $(X^T X)^{-1}$ term; the rank one $(X^T X)^{-1}\Sigma r$ is omitted. Here one sees the effect of the input measurement noise on significance regression: SR will underestimate the variance of $\tilde{r}$ in the $(X^T X)^{-1}\Sigma r$ direction. Under SS, $\Sigma = I$ and $(X^T X)^{-1}\Sigma r \in \mathrm{Range}(W)$ if only if $n_d = n_i$; significance regression naturally avoids the "extra-noisy" direction. However, this is not particular motivating since any given significant vector can be arbitrarily close to $(X^T X)^{-1}\Sigma r$ (that is, have a direction cosine arbitrarily close to unity). Recall that the significance regression objective,

$$\frac{(w^T \tilde{r})^2}{w^T (X^T X)^{-1} w \sigma_e^2}, \tag{8.16}$$

seeks the direction where the ratio of signal $(w^T \tilde{r})$ to variance is greatest. Thus, SR will have a greater preference for the direction where $(X^T X)^{-1}\Sigma r$ has the greatest *relative* impact than SR-MEM would indicate is appropriate. This direction of this bias is

$$\arg \min_{v \in \Re^{n_i}} \frac{(v^T (X^T X)^{-1} \Sigma r)^2}{v^T (X^T X)^{-1} v \sigma_e^2} = \Sigma r. \tag{8.17}$$

This will cause the distribution of $w_1^{opt}(y)$ to be asymmetric about $X^T X r$ — unless $\Sigma = I$. Thus when SS is used with SR, the search for the significant directions is biased towards r relative to using SR-MEM. A basic requirement for a restriction

regressor to be "good" is $r \in \text{Range}(W)$, so the "bias" significance scaling introduces in the search for significance directions is also "good."

8.7 Conclusions

This chapter examined scaling regimes for restriction regressors, and significance regression in particular. The output scalings should be chosen to render the data homoscedastic. Input scalings are a reflection of one's *a priori* beliefs about r. When no such beliefs prevail, existing theory provides only a weak guide. However the theory underpinning significance regression encourages using auto-scaling when the inputs are known without error and using significance scaling when the inputs are corrupted by measurement noise.

Part IV

Process Applications

Chapter 9

Synthetic Examples

9.1 Introduction

The above development focused on theoretical understanding and derivation. This chapter presents numerical studies that illustrate the utility of the new results for problems with multiple outputs and clarify the relationship between the various regressors. The results of the previous chapters for vector output problems, the measurement error model, robust regression, and scaling are illuminated. In the past, some novel estimators have been "proven" using a specific example. Often much of the ensuing debate centered around the "validity" of the chosen example and the "meaning" of the numerical results. In this chapter, the examples are simulation studies using purely synthetic data. The data are not claimed to correspond to any particular "real world" process; rather, the data were generated to conform to the model assumptions and to illustrate the relative effectiveness of various methods for problems that satisfy the model assumptions. The "real world" successes of PLS [62, 63, 75] and the results of the following two chapters are suggested as evidence of the practical utility of SR since the two methods are closely related.

Two measures were employed to evaluate regressor performance. Since the exam-

ples were synthetic, R was known and a point estimate of the trace of the MSE could be computed for each example. The measure was

$$RMS_{\text{MSE}} = \sqrt{\frac{\text{Tr}(\ (\tilde{B} - R)(\tilde{B} - R)^T\ )}{\text{Tr}(RR^T)}}. \tag{9.1}$$

The $\text{Tr}(RR^T)$ term was included to produce a relative error and allow averaging over all one thousand examples.

The second measure was computed based on the PRESS (PRediction Error Sum of Squares). For each example an additional one hundred samples (X_{new}, Y_{new}) were generated from the identical distribution as the training data, but the Y_{new} were not corrupted by error ($E_{new} = 0$). Then

$$RMS_{\text{PRESS}} = \sqrt{\frac{\text{Tr}(\ (X_{new}\tilde{B} - Y_{new})^T(X_{new}\tilde{B} - Y_{new})\ )}{n_s\ n_o}}. \tag{9.2}$$

Since the data were generated with the constraint

$$\sqrt{\frac{\text{Tr}(Y_{new}^T Y_{new})}{n_s\ n_o}} = 1 \tag{9.3}$$

the RMS_{PRESS} was averaged over the examples without normalization. For each simulation, the rank (relative performance) of each regressor was recorded: rank = 1 if no other regressor did better for that example, rank = 2 if one other regressor did better, and rank = 3 if two other regressors did better. If two regressors had the same performance to within 0.1% they were given the same rank. The average rank with respect to both MSE and PRESS was computed.

All examples had ten inputs ($n_i = 10$). The examples for the vector output and robust studies had four outputs ($n_o = 4$), while the MEM and scaling studies used scalar output examples. For each case study, one thousand distinct examples were examined to mitigate sampling effects in the numerical results. Each example was

generated by the appropriate method presented in appendix B. Since both input variances and the values of the regression parameters varied over five orders of magnitude and since there were typically large variances in the input data that had little effect on the output, this exploration shed light on the relative strengths and weaknesses of the four methods for a class of problems that has historically bedeviled OLS. Thirty samples were available for training ($n_s = 30$) for all examples. Where cross-validation was used to determine n_d, ten-way (three-out) cross-validation was employed.

9.2 Vector Outputs

The first suite of simulations examined estimation and prediction for the vector output problems. The regression methods investigated were

- ordinary least squares (OLS, equation 2.4),
- partial least squares using cross-validation (PLScv, appendix A),
- significance regression using cross-validation (SRcv, algorithm 3.1), and
- significance regression using equation 3.21 and the approximate distribution for $(\tau_i^{opt}(y))^2$ defined in section 3.2 (SR).

The examples were generated via the routine described in appendix B.1

The results of the first test, which compared SRcv to PLScv, are shown in Table 9.1. Since PLScv and SRcv are similar, one should not be surprised that the two methods had similar results and outperformed OLS in all measures. As discussed in section 5.4, PLS is *not* optimal for determining the "most significant subspace" of $\Re^{n_i \times n_o}$. This is reflected in the MSE results; the $\overline{RMS_{\text{MSE}}}$ for PLS was almost eight times that of the $\overline{RMS_{\text{MSE}}}$ for SRcv. This difference did not appear to be crucial for prediction; SRcv and PLScv produced almost identical results as measured by $\overline{RMS_{\text{PRESS}}}$.

method	$\overline{RMS_{\text{MSE}}}$	$\overline{\text{rank}}$	$\overline{RMS_{\text{PRESS}}}$	$\overline{\text{rank}}$
OLS	860	3.0	0.36	3.0
PLScv	8.0	1.4	0.19	1.4
SRcv	1.1	1.3	0.19	1.4

Table 9.1. Comparison of PLS and SR using cross-validation over 1,000 examples of synthetic data.

method	$\overline{RMS_{\text{MSE}}}$	$\overline{\text{rank}}$	$\overline{RMS_{\text{PRESS}}}$	$\overline{\text{rank}}$
OLS	860	3.0	0.36	2.8
SR	0.9	1.5	0.27	1.7
SRcv	1.1	1.2	0.19	1.2

Table 9.2. Comparison of the 90 % significance test and cross-validation over 1,000 examples of synthetic data.

method	$\overline{RMS_{\text{MSE}}}$	$\overline{\text{rank}}$	$\overline{RMS_{\text{PRESS}}}$	$\overline{\text{rank}}$
OLS	860	3.0	0.36	2.8
SR	0.9	1.4	0.25	1.7
SRcv	1.1	1.2	0.19	1.2

Table 9.3. Comparison of the "over-simplified" significance test and cross-validation over 1,000 examples of synthetic data.

The second test compared cross-validation to using equation 3.21 with the approximate distribution (F- distribution) developed in section 3.2 operating with a 90% significance criterion. These results are shown in Table 9.2. The two methods used the identical algorithm to compute $\tilde{B}$. SR and SRcv have similar $\overline{RMS_{\text{MSE}}}$, even though SRcv used ten times more computations than SR. Thus, in terms of $\overline{RMS_{\text{MSE}}}$ the significance test performed almost as well as cross-validation and was much less computationally demanding. Interestingly, the $\overline{RMS_{\text{MSE}}}$ for SR reported in Table 9.2 was less than one-eigth of the $\overline{RMS_{\text{MSE}}}$ for PLScv reported in Table 9.1. These numbers are directly comparable since both were generated using the same one-thousand synthetic examples. Next the "over-simplified" significance test of section 3.2 was investigated. This criterion is: reject $\mathcal{H}_0^{2,i}$ if $(\tau_i^{opt}(y))^2 > n_p$. For large n_o and n_s this is a crude approximation to the 50% significance threshold. Table 9.3 shows the results; the "over-simplified" method had similar results to the approximate test with a 90% significance test. For these simulations the results were relatively insensitive to the choice of the significance level. In terms of the $\overline{RMS_{\text{PRESS}}}$, cross-validation was clearly superior to the approximate significance test.

These numerical explorations illustrated several points. For the purpose of prediction partial least squares is virtually identical to the significance regression. However, SR was clearly superior for estimation in these problems. Less computationally demanding alternatives to cross-validation can be developed from the classical viewpoint of significance, but more work is needed on these significance tests. In particular, the relationship between desired objective (*e.g.* $\overline{RMS_{\text{MSE}}}$ or $\overline{RMS_{\text{PRESS}}}$) and choice of significance test needs further work. Still, even the current SR approach using approximate significance test outperformed PLS for estimation while using only one-tenth the computational effort.

9.3 The Measurement Error Model

Next examined were the methods and claims of chapter 6. The regression methods investigated were

- asymptotically unbiased estimation (MEM, equation 6.3),
- ordinary least squares (OLS, equation 2.4),
- the MEM significance regression method (SR-MEM, algorithm 6.1), and
- the classical model significance regression method (SR, algorithm 3.1).

In all cases, a 90% significance threshold was used to evaluate $\mathcal{H}_0^{13,i}$ and to determine n_d. Each example was generated by the method presented in appendix B.2. The examples tended to be collinear in that the singular values of T (the square root of the eigenvalues of $T^T T$) and the values of the regression parameters varied over five orders of magnitude; moreover there were typically large variances in the input data that had little effect on the output. The same one thousand examples were used in all studies. All errors were independent and homoscedastic; however the variances of the errors for each input varied by one order of magnitude. In all examples, thirty samples were used to compute the regressor ($n_s = 30$).

The performance loss due to using $\tilde{r}_{\text{MEM}}$ in place of r for determining V was investigated. Since the examples were synthetic the true r could be used to compute V. In 872 of the 1,000 examples the use of $\tilde{r}_{\text{MEM}}$ in place r affected the RMS by less than 0.1%. Use of r improved the $\overline{RMS_{\text{MSE}}}$ in only 63 of the remaining 128 problems. Use of $\tilde{r}_{\text{MEM}}$ actually lead to a lower $\overline{RMS_{\text{MSE}}}$, but the difference was primarily attributable a single "fluke" example. These synthetic examples support the conjecture that using $\tilde{r}$ in place of r causes only "slight" performance degradation.

Next studied was the effectiveness of SR-MEM; these results are in Table 9.4. SR-MEM had an $\overline{RMS_{\text{MSE}}}$ three orders of magnitude less than that of MEM. Clearly

method	$\overline{RMS_{\text{MSE}}}$	rank
MEM	3,200	2.0
SR − MEM	1.3	1.0

Table 9.4. Comparison of MEM-based methods over 1,000 examples of synthetic data.

method	$\overline{RMS_{\text{MSE}}}$	rank
MEM	3,200	2.0
OLS	120	1.0

Table 9.5. Comparison of asymptotically unbiased estimator versus least-squares estimator over 1,000 examples of synthetic data.

method	$\overline{RMS_{\text{MSE}}}$	rank
MEM	3,200	2.8
SR	5.6	1.5
SR − MEM	1.3	1.3

Table 9.6. Comparison of SR-MEM and SR methods over 1,000 examples of synthetic data.

method	$\overline{RMS_{\text{PRESS}}}$	rank
null estimator	1.00	4.0
OLS	0.145	2.8
SR	0.103	1.6
SR − MEM	0.103	1.6

Table 9.7. MEM prediction performance over 1,000 examples of synthetic data.

SR-MEM mollified much of the difficulty caused by the correlations among the inputs. Next, the "ridging effect" of using $\tilde{r}$ for estimating r was examined; these results are shown in Table 9.5. In this example, the performance degradation due to bias was more than offset by the reduction in variance: OLS reduced the $\overline{RMS_{\text{MSE}}}$ by two orders of magnitude. Last studied was the effectiveness of using SR, which was derived assuming the classical model, with the F-test described in section 3.2. This test is not rigorously correct for MEMs. However this F-test is used here since the purpose of the simulation was to investigate the performance of a method based entirely on the classical model for data generated using a measurement error model. As shown in Table 9.6, SR was almost as good as SR-MEM. Thus, if one does not have knowledge of the error covariances but the data exhibit collinearity these simulations suggest that one can use SR without undue performance loss relative to SR-MEM and with considerable performance benefit relative to the asymptotically unbiased estimator. Although not shown here, similar results were obtained if a 95% or 50% significance criterion was used. Consistent with the results of section 9.2 the performance of SR-MEM and SR did not strongly depend on the choice of significance level for these examples.

One might object that none of the estimators did better than the null estimator: using the estimator "always estimate zero" yields $\overline{RMS_{\text{MSE}}} = 1$. However, most of the synthetic examples had large components of r in directions where T had "small" variance. Thus the "success" of the null estimator was a reflection of the difficult nature of the examples used. With null estimator one disavows using any variance information and instead relies solely on the pre-defined nominal points (or the mean of the training data). With the SR approach the space where estimation is not attempted, and therefore bias may exist, is precisely $\text{Range}(I - W_{n_d} W_{n_d}^T)$. Thus SR proceeds where the data are "significant" and suggests where to be wary. In the subspaces with "significant" input data, the null predictor performed poorly, as shown

by studying the $\overline{RMS_{\text{PRESS}}}$. Since, as shown in claim 6.2, OLS is more appropriate than MEM for prediction problems, the SR-MEM algorithm used

$$V = \frac{1}{n_s}\left(\frac{X^TX}{n_s}\right)^{-1}\left(\Sigma\tilde{r}\tilde{r}^T\Sigma + \sigma_e^2\frac{X^TX}{n_s}\right)\left(\frac{X^TX}{n_s}\right)^{-1} \tag{9.4}$$

and replaced equation 6.15 with

$$\tilde{b} = W_{n_d}(W_{n_d}^TX^TXW_{n_d})^{-1}W_{n_d}^TX^Ty \tag{9.5}$$

when computing $\tilde{b}$ for prediction. The results are shown in Table 9.7. All of the methods did at least six times better than the null-estimator. Indeed, the null estimator was a worse predictor than all other investigated methods for all one-thousand examples. Thus the "superior" estimation performance of the null estimator occurred primarily in the subspace where the data were not "significant" — the space clearly delineated by SR-MEM. Interestingly, SR and SR-MEM had similar predictive performances.

9.4 Data with Outliers

The next suite of simulations examined estimation and prediction for the vector output problems with outliers. The regression methods investigated were

- ordinary least squares (OLS, equation 2.4),
- significance regression (SR, algorithm 3.1),
- M-estimation (M-est, algorithm 7.2), and
- robust significance regression (SR-Robust, algorithm 7.3).

Both robust algorithms used Huber's "proposition 2" objective function; both cross-validation and significance testing were used to choose n_d. In all cases the "over-simplified" test was the significance test used.

The examples for this simulation study could easily have been designed to provide SR-Robust with an overwhelming advantage over SR since the standard SR algorithm has a zero breakdown point. While this would provide some dramatic numbers, little insight would be gained. Instead this study uses data that is only mildly corrupted with outliers: for each error, there was a 5% probability that the error would be drawn form a distribution with three times the standard deviation. According to Huber, "typical 'good data' samples in the physical sciences appear to be well modeled" by this distribution [42, p. 2].

Notice that each sample contains $n_s n_o = 120$ samplings of the error distribution, of which typically only 4 will lie outside the 3σ range of the nominal distribution. Moreover, there are no "extremely gross" errors; this study gives SR the benefit of the doubt and assumes errors of more than an order of magnitude will be detected by experimenter/operator. Thus the outliers in this study are relatively mild and hopefully reasonable for "real" data. The outliers in this study were generated by perturbing the output. However these perturbations are also loosely equivalent to failures of input measurements as well, since the errant input value will tend to lead to misprediction and hence appear as an outlier. The actual examples were generated by the method presented in appendix B.3.

To illustrate the nature of the examples, the first simulation compared OLS to the M-estimator. OLS actually slightly outperformed the M-estimator: the dominant feature of these examples is collinearity, not outliers. The outliers did cause difficulty; even though the examples were generated with the same parameters, the $\overline{RMS_{\text{PRESS}}}$ degraded from 0.36 (as reported in Table 9.1) to 0.42. These numbers can only be compared loosely because the examples were not identical for the two simulations.

Examining Table 9.9 reveals the efficacy of the robust method. SR, which (non-robustly) treats the collinearity, brings the $\overline{RMS_{\text{PRESS}}}$ down to 0.28, while SR-Robust, which addresses the (mild) outliers, has a $\overline{RMS_{\text{PRESS}}}$ of 0.22. The robust method was a consistently better regressor as revealed by the ranks.

As revealed in section 9.2, the significance test is useful for estimation, but cross-validation is a consistently better method for choosing n_d for prediction purposes. For the SR-Robust, cross-validation was performed on the "robust" PRESS: the contribution of each output was weighted by it's M-estimation weight, thereby diluting the effect of outliers. The results are shown in Table 9.10. Cross-validation does provide the SR method with a measure of robustness: the cross-validated SR had the same $\overline{RMS_{\text{PRESS}}}$ as SR-Robust using the "over simplified" significance test. The same examples were used in the two different simulations, so the numbers are directly comparable. However, the "robust" cross-validation gave the best performance, as measured by both the $\overline{RMS_{\text{PRESS}}}$ and the ranks.

9.5 Scaling

The final set of simulations examined scaling issues and the recommendations of chapter 8. First the two most popular scaling methods, auto-scaling and significance scaling, were applied to the identical one thousand MEM examples used in the previous section. The regression methods investigated were

- ordinary least squares (OLS),
- significance regression (SR, algorithm 3.1),
- significance regression with auto-scaled data (SR-AS), and
- significance regression with significance scaled data (SR-SS).

Both significance testing using the "over-simplified" significance test and cross-validation using the PRESS were evaluated; the results are shown in Tables 9.11 and 9.12 respectively. For both of the approaches for determining n_d and all four performance measures, significance scaling consistently and dramatically outperformed auto-scaling. However, the best performance was achieved via "no scaling." Not scaling slightly but consistently outperformed significance scaling, as shown by the very similar $\overline{RMS}$ values but markedly different ranks. These results were consistent with the suggestions given in section 8.6: significance scaling is better than auto-scaling for mitigating the effects of input measurement errors when one uses SR, but knowing the true *a priori* distribution for r is best of all. The "no scaling" approach correctly reflected the true *a priori* distribution for r, since all directions were equally probably *a priori*. In general, however, one does not have such knowledge, so significance scaling would be preferred.

The final simulation investigation examined scaling for the situation where the theory was least adequate: heteroscedastic input measurement errors. Each example was generated by the method presented in appendix B.4. These examples were the same as the examples of the prior section, except that the first ten examples had three times the input noise standard deviation, while the last ten examples had one third the input noise standard deviation. The prior simulation studies verified specific claims based theoretical analysis, but this simulation had no claims to verify. Therefore, this study should be considered a weak illustration of the relative strengths of the method for these specific examples. Given that these simulations are admittedly contrived, one should be careful extrapolating these heteroscedastic results to "real" regression problems.

The regression methods examined were

- significance regression (SR, algorithm 3.1),

method	$\overline{RMS_{\text{MSE}}}$	rank	$\overline{RMS_{\text{PRESS}}}$	rank
OLS	1000	1.4	0.42	1.4
M – est	1000	1.5	0.43	1.6

Table 9.8. Comparison of OLS and M-est over 1,000 examples of synthetic data.

method	$\overline{RMS_{\text{MSE}}}$	rank	$\overline{RMS_{\text{PRESS}}}$	rank
OLS	1000	3.0	0.42	2.8
SR	1.0	1.6	0.28	1.7
SR – Robust	1.0	1.3	0.22	1.4

Table 9.9. Comparison of SR-Robust and SR using significance testing over 1,000 examples of synthetic data with outliers.

method	$\overline{RMS_{\text{MSE}}}$	rank	$\overline{RMS_{\text{PRESS}}}$	rank
OLS	1000	3.0	0.42	2.8
SR	1.3	1.5	0.22	1.6
SR – Robust	1.3	1.3	0.21	1.4

Table 9.10. Comparison of SR and SR-Robust using cross-validation over 1,000 examples of synthetic data with outliers. SR-Robust used "robust" PRESS.

method	$\overline{RMS_{\text{MSE}}}$	rank	$\overline{RMS_{\text{PRESS}}}$	rank
OLS	120	3.8	0.15	3.5
SR – AS	44	2.9	0.12	2.7
SR – SS	3	1.9	0.10	2.0
SR	3	1.4	0.10	1.7

Table 9.11. Comparison of scaling regimes for treating input measurement errors when using an "over-simplified" test over 1,000 examples of synthetic data.

- significance regression using the *ad hoc* scaling of equation 8.14 (SR-scaled), and
- robust significance regression using Huber's "proposition 2" objective function (SR-Robust, algorithm 7.3).

All three methods used the "over-simplified" significance test. The results are shown in Table 9.13. The worst method was the *ad hoc* scaling, which did worse than unscaled SR. The most successful method was robust significance regression.

9.6 Conclusions

This chapter investigated regression issues via simulation. The synthetic examples were collinear by design and had well defined error distributions. PLS was shown to be a suboptimal estimator for vector output problems, although the difference was not significant for prediction. Significance testing was successful for estimation purposes, but cross-validation on the PRESS was clearly superior for prediction. The SR-MEM procedure was effective for the measurement error model; however, the "classical" SR method, which does not explicitly account for input measurement errors, was almost as effective. SR-Robust was a notable success; even for problems with mild outliers, SR-Robust delivered superior results. When using SR with data known to have measurement errors, significance scaling is the clear choice over auto-scaling if the one does not possess an *a priori* distribution for r. If these input measurement errors are also heteroscedastic, SR-Robust appears to be a useful method, although more work is needed in this area.

method	$\overline{RMS_{\text{MSE}}}$	rank	$\overline{RMS_{\text{PRESS}}}$	rank
OLS	120	3.7	0.15	3.5
SR – AS	49	2.8	0.12	2.6
SR – SS	9	1.9	0.11	1.9
SR	10	1.5	0.10	1.7

Table 9.12. Comparison of scaling regimes for treating input measurement errors when using cross-validation over $1,000$ examples of synthetic data.

method	$\overline{RMS_{\text{MSE}}}$	rank	$\overline{RMS_{\text{PRESS}}}$	rank
SR – scaled	62	2.7	0.13	2.5
SR	16	1.6	0.11	1.7
SR – Robust	8	1.4	0.11	1.7

Table 9.13. Comparison of the *ad hoc* heteroscedastic methods over $1,000$ examples of synthetic data.

Chapter 10

Composition Inference for Control of a Packed Bed Reactor

10.1 Introduction

In many industrial processes the variables to be controlled cannot be measured fast enough to achieve desired control objectives or, in some cases, cannot be measured at all. However, readily available process measurements such as temperatures, pressures, and flowrates contain much process information. Exploiting relationships between these secondary variables and the variables to be controlled (primary variables) holds the potential to make more demanding control objectives accessible. For instance inferring the primary variables from faster secondary measurements can improve the performance of the closed loop system. This chapter focuses on the application of robust inferential control schemes to a packed bed methanation reactor. Before progressing, one potential source of confusion needs to be addressed. This investigation focuses on developing control-relevant predictions of primary (controlled) variables from the secondary variables. In the control literature, this problem is usually described as an *estimation* problem: one desires to estimate an unknown value. However,

since this problem does not address the problem of identifying the parameters of an underlying "true" model, but rather focuses on using various input values to predict various output values, this chapter will deviate from the control literature, maintain consistency with the development in this thesis, and refer to the *prediction* problem.

Two different design approaches for inferential control wil be compared. The first approach, which can be derived from the Kalman filter, assumes that the system can be modeled by a linear superposition of the effects of disturbances and manipulated variables on the primary and the secondary variables. This approach was initially proposed by Brosilow [5] and thus will be called the "Brosilow-type inferential scheme." The second approach uses restriction regression. When the work was conducted, the significance regression framework had not yet been developed, so the scheme was in fact developed using the then state-of-the-art PLS algorithm described in appendix A. However, since this prediction problem had scalar output, the result was in fact identical to using significance regression with cross-validation on the PRESSS. Hence, the second approach is called the "SR scheme." The SR scheme, in contrast with Brosilow's scheme, does not assume that of the effects of disturbances and manipulated variables on the primary and secondary variables obey superposition. In this sense, the SR scheme rests on fewer linearity assumptions the Brosilow scheme.

The three main issues for comparing these schemes are:

- *Accuracy of the inferential predictor.* The goal of the inferential predictor is to accurately predict the controlled variables. Due to the nonlinearity of the process under study, the SR technique will be shown to result in significant performance improvement compared to the Brosilow approach.

- *Robustness of the inferential scheme.* Robustness is essential due to the nonlinearity of the process and the wide range of operating conditions considered. Here robustness means "operability in the presence of model uncertainty." In

this work robustness properties are formally investigated using Structured Singular Value (SSV) Theory [13]. The SSV framework represents the uncertainty as norm-bounded perturbations to the transfer functions.

- *Appropriateness of uncertainty descriptions.* When using SSV theory, a key issue is the selection of an appropriate (*i.e.*, not overly conservative) uncertainty representation. Some problems encountered while using the Brosilow scheme for the reactor (and discussed below) illustrated the difficulties caused by an overly conservative uncertainty description. Briefly, overly conservative uncertainty descriptions lead to overly conservative controllers that sacrifice performance to maintain robustness to unrealistic perturbations.

The physical system studied was a packed bed methanation reactor yielding partial conversion of reactants. Webb [95] described the experimental system. The specific reaction studied is the methanation of CO_2:

$$CO_2 + 4H_2 \rightarrow CH_4 + H_2O. \tag{10.1}$$

The problem was to control the maximal bed temperature and the exit concentration of the reactants by manipulating the recycle flowrate and the heater power at the inlet to the reactor. The variables had to be controlled in the presence of changes in the reactor wall temperature. The control objectives were to maintain stability and to obtain good performance over a range of different operating conditions.

Several prior investigations have examined implementing inferential control using Kalman filter techniques [5, 27, 47, 52]. These efforts placed special emphasis on measurement selection and can be grouped into two main categories delineated by the uncertainty description. In the first approach [27, 47, 52], the system was modeled using the partial differential equations of heat and mass balance. The controller design problem was addressed from within the Kalman filter framework and model

uncertainty was described with additive Gaussian noise in the system differential equations. The second approach to the measurement selection problem describes model uncertainty as norm bounded perturbations to the transfer functions describing the system [5, 54]. The mathematical artifice of additive Gaussian noise is generally considered a less realistic description of model uncertainty than the mathematical artifice of norm-bound perturbations to the transfer functions. Moreover, the norm-bounded uncertainty is relatively easy to quantify from experimental data [96] and permits facile treatment of robustness issues[66]. The norm-bounded uncertainty representation was initially introduced in the context of measurement selection by Brosilow [5]. The application of this technique to the fixed reactor is presented in detail in the work of Webb and Budman [96, 97].

As mentioned above, the objective is to design an inferential scheme which will result in good performance in a range of different operating conditions. The system under study, the fixed bed reactor, is known to exhibit highly nonlinear behavior in the selected window of operation. In all the aforementioned studies the predictor was designed using a linearization of the system around a preselected nominal operating condition. Therefore, inaccurate prediction is expected when the system is operated "far" from this nominal condition.

Significance regression (SR) uses a different approach to deal with the nonlinearity of the process. The SR approach establishes a linear relation between the primary and both the secondary and the manipulated variables over a wide range of operating conditions. This has two direct benefits. First, this technique requires weaker linearity assumptions than the Kalman filter technique. The Kalman filter method assumes that the effects of disturbances and manipulated variables on both the primary and secondary variables obey superposition. In the SR approach these assumptions are not required. Second, since the predictor is built using data from a range of operating conditions, changing from one operating condition to another does not require the

development of a new predictor. From the outset, the need to function under different conditions is addressed. A drawback of the SR technique is that SR has not been extensively studied for dynamic modeling. In this work the dynamics of the predictor were corrected by augmenting the static predictor with a lead-lag compensator. Using the SR technique the accuracy of the predictor tends to increase as the number of measurements is increased. According to Brosilow increasing the number of sensors may result in increasing sensitivity to model error. However, results below show that this increased sensitivity is an artifact of the assumed uncertainty structure.

This study did not address the full measurement selection problem, since only a fixed number of measurements was available at fixed positions along the axis of the reactor. One was not free to either increase their number or change their position. To keep the problem manageable, this study only examined disturbances in the reactor wall temperature. The steady-state bed temperatures and the exit concentration change significantly with small changes in the wall temperature.

10.2 Brosilow's Inferential Control

The Brosilow inferential control scheme was developed for systems that can be described by:

$$\begin{aligned} s &= G_{sd}d + G_{sm}m \\ c &= G_{cd}d + G_{cm}m \end{aligned} \tag{10.2}$$

where s is the vector of secondary measurements, c is the vector of controlled variables, m is the vector of manipulated variables, d is the vector of unmeasured disturbances, and G_{ij} is the transfer function relating output vector i to input vector j.

In principle, all possible combinations of temperature (secondary) measurements

should be considered when selecting the measurement set for prediction. In the investigation of the Brosilow method, the search was limited to a *single* "best measurement." Beyond the obvious motivation to keep the predictor as simple as possible,

- changes in reactor wall temperature were the only type of disturbance considered in the experiments. If the sensor noise is negligible, accurate predictions of a single disturbance are possible using one single measurement. Intuitively, one would expect that using more measurements would lead to greater prediction accuracy; however, for the nonlinear system under consideration, this is not necessarily the case [8]. Moreover,

- the robust stability criterion is more difficult to satisfy when additional measurements are used for prediction. To illustrate this point the case where two secondary variables are considered for prediction was investigated. Twelve different pairings were examined, each one including the maximal bed temperature and one of the 12 measurements along the reactor. A necessary condition for robust stability, based on steady-state conditions, is developed in [8]. None of the candidate pairs considered for prediction satisfied this condition.

Following these considerations, one single sensor was used to infer the effect of the external disturbance on both controlled variables: maximal bed temperature and exit concentration. Based on the method of Lee and Morari [54] the "best sensor" was the one closest to the hotspot but never crossed by it. This thermocouple was positioned at approximately 5 inches away from the reactor exit. The controller was designed based on the IMC factorization of Holt and Morari [37]. This factorization procedure led to a dynamically decoupled system with minimal possible delays in the controlled variables. The IMC robustness filter time constant is found from robust stability requirement. Based on a Structured Singular Value test this time constant was found to be 83 seconds. The controller design was presented in greater detail

in Webb *et al.*[97]. Summarizing, the two main deficiencies observed when using the Brosilow scheme for the reactor were

- Inaccurate prediction and
- High sensitivity to model uncertainty.

The objective of the predictor was to predict accurately the primary variables (exit reactant concentration and maximal bed temperature) over a range of different operating conditions. To partially test if this objective was achieved when using the predictor proposed above, the reactor was operated in open loop around eighteen different operating points in the selected window of operation. At each one of the operating points the temperature values obtained with the 12 thermocouples and the gas chromatographic readings of the concentration were recorded when the system reached steady-state.

The accuracy of Brosilow's predictor was tested. This predictor used sensor 13 for prediction since this measurement was determined in the previous section to give the best steady-state performance. For simplicity the comparison was conducted for the exit reactant concentration only. Using the temperature value from thermocouple 13 and the corresponding value of the manipulated variables, the Brosilow inferential predictor for concentration is:

$$y = (G_{cm} - G_{cd}G_{sd}^{-1}G_{sm})\delta m + G_{cd}G_{sd}^{-1}\delta T + y_n \tag{10.3}$$

where δm is the deviation in manipulated variables, δT is the temperature deviation at sensor 13, and y_n is the concentration at the nominal operating condition.

The development of this predictor is given in Budman *et al.*[8]. From the results shown in Table 10.1 one observes that errors up to 0.65% vol of CO_2 occurred. The differences between the predictions and the actual values tended to increase as the system was operated "far" from the selected nominal condition of 17 slpm recycle

Type of error	Brosilow 1 sensor (% vol)	Brosilow 12 sensors (% vol)	SR 1 sensor (% vol)	SR 12 sensors (% vol)
worst error	0.65	0.77	0.19	0.06
RMS error	0.20	0.23	0.08	0.02

Table 10.1. Steady-state experiments using Brosilow's and SR predictors

flowrate and 30% heating power.

As mentioned in section 10.1, Brosilow concluded that for his choice of sensor noise and process noise respectively, a larger number of measurements will result in more accurate prediction. But does this conclusion still apply when the reactor is operated at different operating points? To test this question, a static predictor which uses all available 12 temperatures along the reactor was computed. Table 10.1 compares gas chromatographic readings for the different conditions with concentration predictions computed from equation 10.3. From the results one can conclude that overall the accuracy decreased; both the RMS and worst errors increased as compared to the one measurement prediction case. However, certain improvements were observed for points "close" to the nominal operating point. Since Brosilow's predictor was designed for linear systems with additive Gaussian process and sensor noise, the pattern of improvement "close" to one operating point coupled with performance deterioration "far away" should not be surprising. The effect of noise terms on the predictions can be expected to be filtered by this predictor as more measurements are used. However, using more measurements will not necessarily make the predictor more accurate when applied to a nonlinear system operated at different conditions; the Brosilow predictor linear superposition assumption may break down. This linearity assumption can be arbitrarily inaccurate for nonlinear systems as demonstrated by the results in this section. This shortcoming motivated the search for an alternative predictor.

In the Brosilow inferential scheme the controller is the inverse of the plant G_{cm} at

steady-state, even though the condition number of G_{cm} was 812. This large condition number resulted from the controlled variables being highly correlated in parts of the operating window. This difficulty was found both experimentally in the present work and through numerical simulations in a previous work by Mandler[60]. In general controllers for plants with "large" condition number are sensitive to model uncertainty [66], partially explaining why the robust stability condition could not be satisfied even for a small number of measurements. An additional difficulty resulting from the high condition number was that the manipulated variables are easily driven to their saturation limits for relatively small external disturbances. For example, the heating power reaches its saturation limits for a change of the order of 3 degrees in the wall temperature. Therefore an alternative dynamic compensator was sought to avoid the aforementioned problems.

10.3 SR Prediction

The prior section examined the Brosilow scheme and found several problems, including an alarming number of linearity assumptions and undue sensitivity to modeling errors. This section develops a novel approach to inferential control that rests on restriction regression and the SSV. The resulting scheme requires fewer linearity assumptions and tends to be more robust.

10.3.1 SR Predictor

The SR design procedure seeks to minimize the squared error between measurements of primary variables and predictions of these primary variables made from secondary measurements for a set of different operating conditions. For this application, the primary variables were outlet CO_2 concentration, while the secondary measurements were both manipulated variables and all twelve temperature measurements along

the reactor. One should note that because only steady-state data were collected, a static predictor was obtained. A simple parabolic interpolation of the three highest contiguous temperature measurements along the reactor yielded good predictions of the maximal temperature. Therefore, the SR regressor was used solely to predict the outlet concentration while the parabolic interpolation was used to predict the temperature. In general one expects that the accuracy of the regression will improve as the number of measurements is increased; hence, all 12 fixed thermocouples were used for prediction.

In the SR technique the regression was performed directly from manipulated variables and temperature measurements to the concentration. This was the main reason that the SR predictor was expected to produce more accurate predictions than Brosilow's predictor. For Brosilow's (and Kalman filtering) schemes the transfer functions G_{sm}, G_{sd}, G_{cd} and G_{cm} are separately identified and then the concentration was computed from equation 10.3 (or a variant of equation 10.3 in the case of the Kalman filter). Thus, the coefficients for δm and δT in the predictor are assembled from algebraic operations of the separately identified transfer functions, in contrast to the SR predictor where these coefficients are computed directly from the regression procedure. The more direct SR approach leads to improved prediction accuracy, as will be shown below.

For the reactor data, fourteen data points representing fourteen different steady-state operating conditions were available. The data are tabulated in appendix C. Each data point consists of twelve temperature measurements and 2 manipulated variables; thus, the problem would appear to have an exact solution. However, the input data are highly correlated, making the problem very ill-conditioned. SR was used to address this ill-conditionedness (collinearity). The scaling was done in two steps on a purely *ad hoc* basis. First the temperatures were significance scaled. The noise level of the two manipulated variables was not known, so after the the tem-

peratures had been significance scaled the entire data set was auto-scaled. Since the primary function of the SR predictor was to predict the compositions, cross-validation was used to determine $n_d = 7$. Further details on the development of the SR predictor are in [8].

The steady-state accuracy of this SR predictor was tested using the same data used for evaluating the Brosilow scheme; Table 10.1 contains the results. The worst error between the SR prediction and the actual concentration was 0.06% vol in CO_2 concentration. This was a considerable improvement in the accuracy over the Brosilow type predictor. The conjecture that the accuracy of the SR predictor may increase with an increase in the number of measurements was also tested. A SR predictor was computed which uses one single temperature for prediction; sensor 13 was selected since this was the sensor selected for Brosilow's predictor. This single temperature SR predictor was compared to the prediction obtained with the 12 sensor based predictor over all 18 operating points. The worst error obtained was 0.19% CO_2 concentration, that is, 3 times larger than the worst error obtained when all 12 temperature measurements were used for prediction.

Since this SR predictor was designed using only steady-state information, one would expect the predictor to be accurate at steady state but inaccurate for dynamic situations. To test the dynamic accuracy of the predictor the time response of the actual exit concentration to a step in recycle flowrate was measured (see Figure 10.1). The time constant for this response was roughly 1200 seconds. A time constant of approximately 300 seconds was observed (see Figure 10.1) when concentration was computed with the SR static predictor, clearly indicating that the predictor dynamics obtained using the steady-state data were inaccurate.

To correct this situation the SR predictor was augmented with a simple lead-lag compensator:

$$\delta y_o = \frac{1 + \tau_A s}{1 + \tau_B s} \delta y_o^{SR}. \tag{10.4}$$

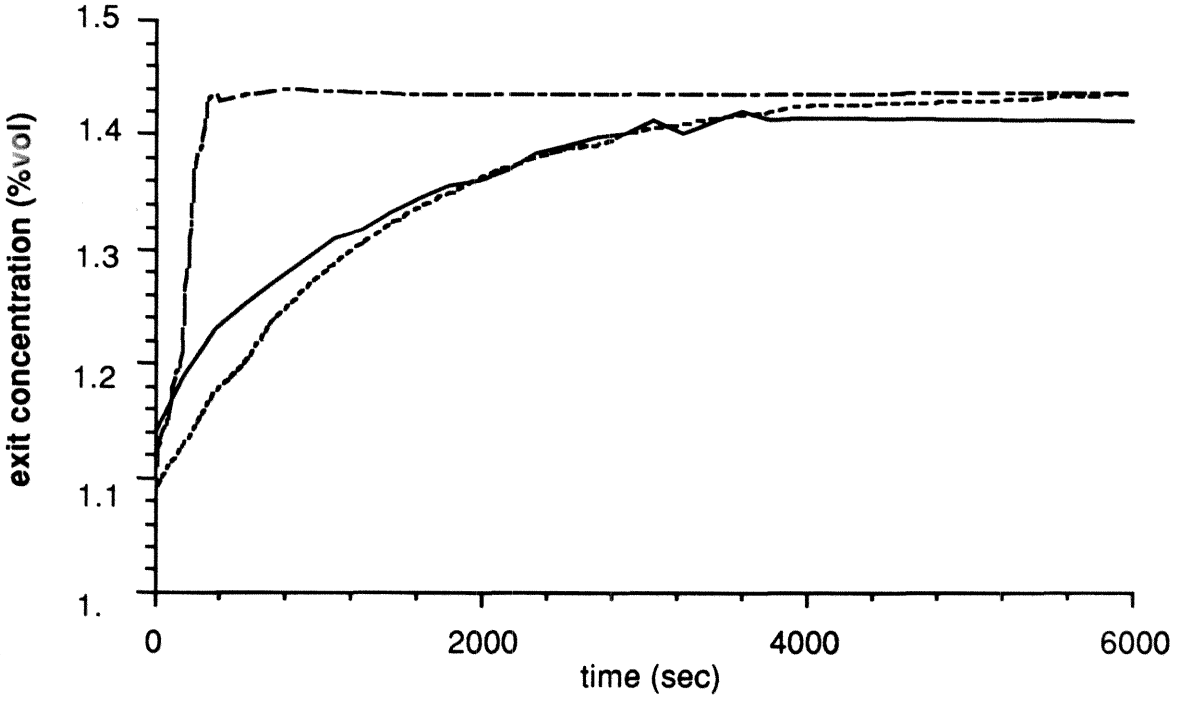

Figure 10.1. The time response of concentration to step in recycle flowrate for concentration: solid line – measured with GC; dot-long-dash line – predicted with SR predictor based on steady- state data; and dotted line — predicted concentration with SR predictor augmented by lead-lag compensator.

The zero and pole of this compensator, τ_A and τ_B respectively, were obtained by fitting the step response to recycle flowrate shown in Figure 10.1 around the nominal operating condition(17 slpm recycle and 30% heating power). The time response for the concentration using the predictor given by equation 10.4 is shown in Figure 10.1.

The predictor given by equation 10.4 has the same structure as the suboptimal predictor proposed by Brosilow as an alternative to the relatively complex optimal Kalman filter. Brosilow originally applied his inferential control to a distillation process. For that process the differences between the closed loop performance of optimal Kalman filter and the suboptimal predictor were minimal.

10.3.2 Uncertainty Description

The SR predictor proposed in this section uses a large number of measurements for prediction. This raises the question of an appropriate uncertainty description. Brosilow's analysis assumed that uncertainty was associated with each one of the measurements. Using this uncertainty description in the SR scheme would be extremely conservative for the analysis, since the uncertainty regions for all the measurements would be added together. The worst possible combination of all these individual uncertainties would then define the uncertainty region for the predictor. Since the SR scheme assumes that correlations between the secondary variables exist, such an uncertainty description would be inherently conservative and would become more conservative as more measurements are considered. Therefore, an alternative uncertainty description is proposed which lumps together the individual uncertainty bounds into a more compact structure.

The key insight motivating to the lumped uncertainty description is realizing that the secondary measurements are only a means to an end. Uncertainties in these variables were only important inasmuch as they affect the overall accuracy of the predictor. Thus, only the uncertainty of the predictor was considered; the identification of the individual uncertainties was completely bypassed. The first step was to identify a transfer function which relates the manipulated variables directly to the predicted controlled variables. To identify this transfer function, the SR predictor for concentration and the parabolic interpolation algorithm for the maximal temperature were implemented on the experimental system. Subsequently linear transfer functions relating the predicted controlled variables to the two manipulated variables and the disturbance are directly identified from step experiments. The resulting block diagram is shown in Figure 2. The actual controlled variables are represented by c

	Gain	Time constant(sec)	Delay(sec)
G_{YH}	-0.007 - -0.015	1800-2200	0-160
G_{YR}	0.06 - 0.16	1193-1507	0
$G_{T_{max}H}$	0.12 - 0.3	500-1000	625-750
$G_{T_{max}R}$	-3.0 - -2.75	245-255	0

Table 10.2. Uncertainty bounds for gains, time constants and delays for the transfer functions relating manipulated variables to the predicted concentration using SR and the interpolated maximal bed temperature.

and the predicted variables by c'. The analysis was simplified by assuming

$$\hat{G}_{cm} = \hat{G}_{c'm} \tag{10.5}$$

The procedure for identifying the nominal transfer functions and the respective uncertainty elements was the same as the one used in previous work [96]. The main idea was to model the system through step response experiments around different operating conditions in the window of operation. As before, the system can be approximated by a first order system with a delay. The nominal models for $G_{c'm}$ and the variation in the time constants and delays along the window of operation are identified from step responses around different operating points(see Table 10.2). This parameter variation can be translated using the technique of Laughlin [53] to an uncertainty description suitable for robustness analysis.

The uncertainty structure in Figure 10.2 is defined by:

$$G_{c'm} = \hat{G}_{c'm} + \Delta^*_{c'm} \tag{10.6}$$

where

$$\hat{G}_{c'm} = \begin{pmatrix} \hat{G}_{YR} & \hat{G}_{YH} \\ \hat{G}_{T_{max}R} & \hat{G}_{T_{max}H} \end{pmatrix} \tag{10.7}$$

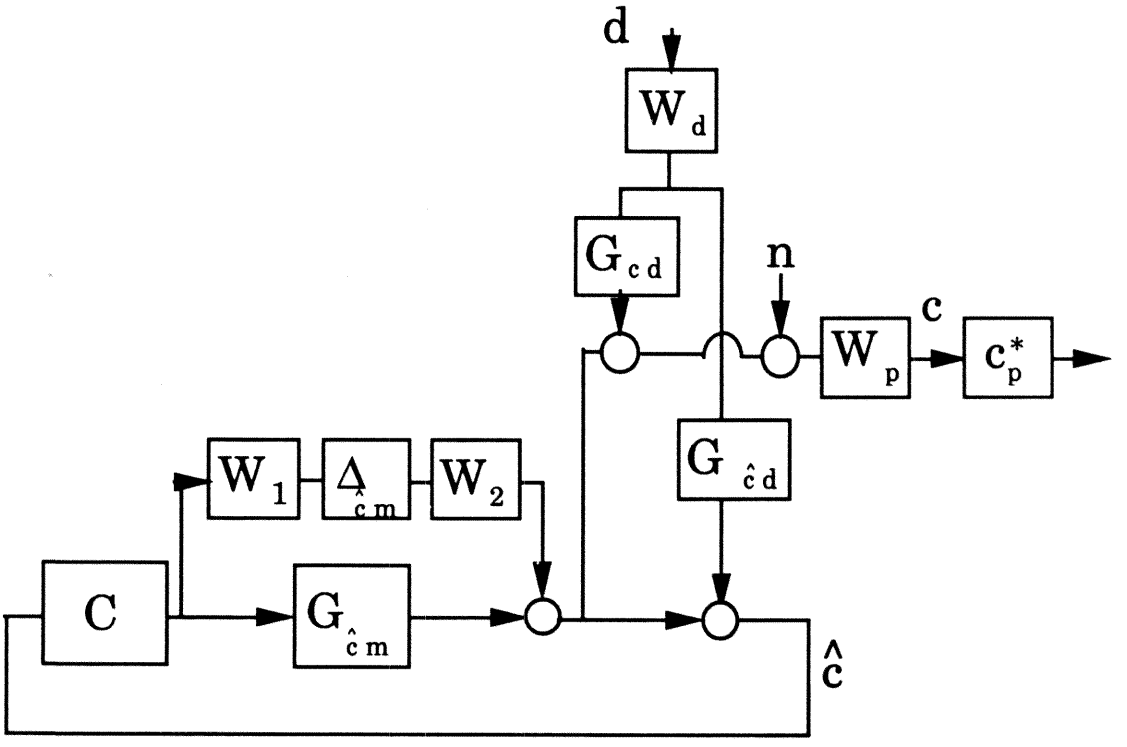

Figure 10.2. Block structure for steady-state performance analysis for the SR inferential scheme

$$\Delta^*_{c'm} = W_{e_2}\Delta_{c'm}W_{e_1} \tag{10.8}$$

$$W_{e_1} = \begin{pmatrix} W_{YR} & 0 \\ W_{T_{max}R} & 0 \\ 0 & W_{YH} \\ 0 & W_{T_{max}H} \end{pmatrix} \tag{10.9}$$

$$W_{e_2} = \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{pmatrix} \tag{10.10}$$

$$\Delta_{c'm} = \mathrm{diag}(\Delta_1, \Delta_2, \Delta_3, \Delta_4) \tag{10.11}$$

$$\bar{\sigma}(\Delta_i) \leq 1 \quad i = 1,4. \tag{10.12}$$

T_{max} is the maximal bed temperature and Y is the exit concentration.

The actual value of the maximal temperature was identical to its prediction since

these quantities are obtained using the same parabolic interpolation with three temperature measurements. Therefore c differs from c' only in the value of the concentration. The predicted value was obtained with the SR regression while the real value was measured using the gas chromatograph. For the purpose of the analysis, an additive noise term n was added to the real concentration to account for the difference between the predicted and the real concentration due to regression errors. The magnitude of this noise was selected to be approximately equal to the worst error between predictions and real values according to the results in Table 10.1 , *i.e.,* $\pm 0.06\%$ CO_2 concentration.

10.3.3 Design of the Dynamic Compensator

As discussed above, ill-conditioned plants may be difficult to control and may be particularly difficult to control integrally (drive all errors to zero) [8]. Since the reactor exhibits a large condition number, a pragmatic decision was made regarding the performance requirements for the control scheme. Accurate control of the concentration requires integral action. However, if one only wishes the temperature to be maintained at a reasonable level to avoid catalyst damage, then offset can be tolerated in the maximal bed temperature. Therefore integral action was specified only for the concentration. Since the concentration reacts faster to recycle flowrate than to heating power, the concentration was paired to the recycle flow, and the maximal temperature was paired to the heating power. This pairing was expected to give better performance for concentration control than the reverse pairing. Explicit conditions have been developed for the pairing problem when integral action is required for all loops[66]. However these conditions do not apply to this problem since integral action is not required for the maximal temperature. Development of a more formal approach to the pairing selection than the engineering heuristic above is left for future investigations.

A simple diagonal controller was used with proportional-integral control for the concentration and only proportional control for the maximal bed temperature:

$$C = \begin{pmatrix} K_1(1 + \frac{1}{\tau_1 s}) & 0 \\ 0 & K_2 \end{pmatrix}. \tag{10.13}$$

The expected offset between measured and predicted concentration, expressed by the noise n, was due to the regression error and was not affected by the selection of the controller parameters. On the other hand, the controller parameters can be selected to minimize the expected offset in the maximal temperature. The procedure used by Lee and Morari [54] for measurement selection was used for selection of the controller parameters. For analysis the controlled variables c and the disturbance d are multiplied by a performance weight W_p and a disturbance weight W_d respectively.

The expected variation for the maximal temperature was $\pm 15°C$ [8]. To scale the variation in maximal temperature and to obtain integral action in concentration the performance weight W_p was

$$W_p = \begin{pmatrix} \infty & 0 \\ 0 & \frac{1}{15} \end{pmatrix}. \tag{10.14}$$

The disturbances weight is $W_d = 2$, representing disturbances of $\pm\ 2°C$. The block diagram with the weights is shown in Figure 10.2. This block diagram can transformed into the M$-\Delta$ structure, from which the controller parameters can be computed from a two-step iterative procedure as described in [8].

10.4 Closed-Loop Experiments

The main goal of the closed loop experiments was to test the stability and performance of the two inferential control schemes. By comparing the theoretical predictions of stability and performance to the experimental results, one can verify the identified

uncertainty of the process. The stability and the performance results are presented separately for each of the two schemes.

10.4.1 Experiments for Brosilow's Scheme

The experiments described below are conducted using thermocouple #13 for inference. Both controlled variables, maximal temperature and exit concentration, were inferred from sensor #13.

Stability Results. As shown in [96], the stability of the closed loop depends only on the uncertainty in G_{sm}. Therefore, one can validate the uncertainty Δ_{sm} by comparing the experimental results to the theoretical stability results. The IMC time constant required for robust stability as computed in Budman [8] is $\lambda \geq 82$sec.

In order to check this result in the experimental system, the disturbance rejection problem was studied for different operating conditions using different filter time constants: 58 sec, 82 sec and 140 sec. These experiments were conducted in the following manner:

1. The pressure in the Dowtherm container was increased resulting in a slow increase in wall temperature.

2. The controller was activated at a predefined operating point.

3. The pressure inside Dowtherm container was released causing a step disturbance in wall temperature.

4. The manipulated and controlled variables are monitored and stability of the closed-loop determined.

The operating conditions for which the experiments were performed and the stability results are presented in Table 10.3. As is apparent from experiments 3, 4, 5 and 6 in Table 10.3, stabilizing the system for lower recycle flow rates was more difficult.

Exp#	Initial Operating Conditions	Robustness Filter Time Constant	Magnitude of Disturbance	Stability Results
1	15% Power 13 slpm Recycle 256°C Wall Temp.	140 sec	-1°C	Stable
2	15% Power 13 slpm Recycle 256°C Wall Temp.	82 sec	-1°C	Unstable
3	15% Power 13 slpm Recycle 256°C Wall Temp.	58 sec	-1°C	Unstable
4	15% Power 15 slpm Recycle 256°C Wall Temp.	58 sec	-1°C	Stable
5	5% Power 17 slpm Recycle 257°C Wall Temp.	58 sec	-2°C	Stable
6	15% Power 19 slpm Recycle 256°C Wall Temp.	58 sec	-2°C	Stable

Table 10.3. Closed-loop inferential control experiments

That is, for small recycle flowrates the system appears to be most sensitive to model mismatch. By increasing the filter time constant, stability was maintained at low flow rates, but at the expense of a significantly reduced speed of response (experiment 1).

The system uncertainty was identified for flow rates between 12 slpm to 20 slpm. However using the filter time constant computed from the robust stability condition (82 sec), instability was obtained for an operating condition corresponding to 13 slpm recycle flowrate and 15% heating power (experiment 2). This was not surprising considering that during the closed loop experiment the recycle flow rate becomes less than the minimal value of 12 slpm for which the uncertainty was identified. This situation is clearly illustrated in Figure 10.3. Therefore as long the manipulated variables during the experiment remain inside the preselected window of operation, the stability condition obtained through μ analysis was verified.

Performance Results. The time history of the controlled variables associated

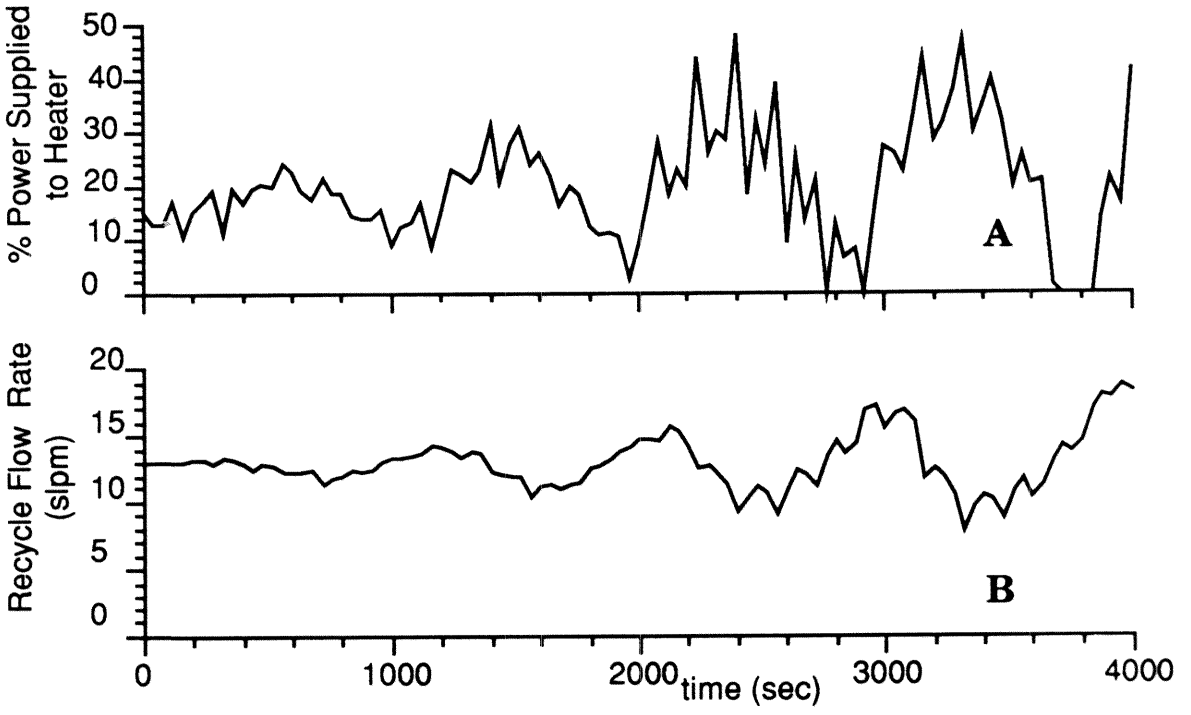

Figure 10.3. Response of the manipulated variable to step decrease in wall temperature (experiment 2).

with Experiment 5 are presented in Figure 10.4. Also shown in this figure are the setpoints and the expected deviations for open-loop conditions. As seen in these figures, the inferential control scheme rejects the disturbance in wall temperature but leaves significant steady-state offsets in both controlled variables. For simplicity, only the steady-state performance criterion given, developed in [8], is compared to the experimental results. More specifically,

$$\frac{\| W_p c(0) \|_2}{\| W_d^{-1} d(0) \|_2} \leq \frac{1}{c_p^*}, \tag{10.15}$$

where $1/c_p^* = 0.42$ for thermocouple #13, was evaluated for all the experiments. Since the inequality was satisfied for all the performed experiments, one sees that the predicted steady-state uncertainty was not invalidated.

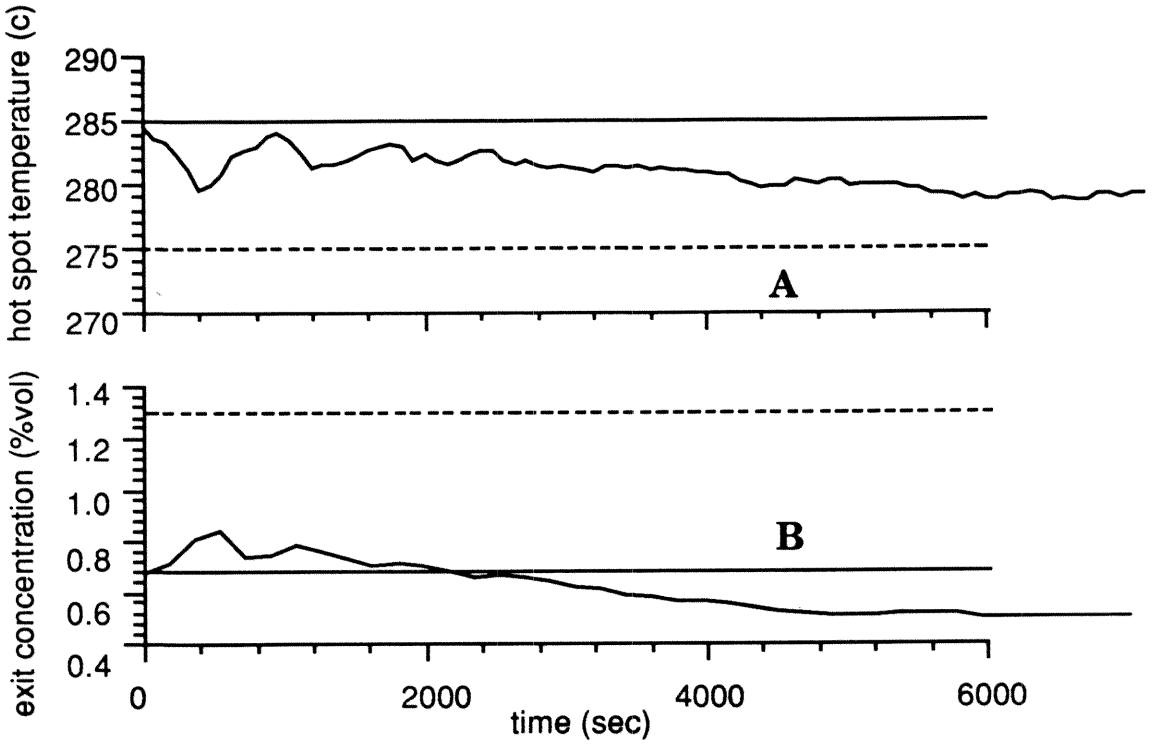

Figure 10.4. Response of controlled variables to step decrease in wall temperature (experiment 5), using Brosilow's inferential scheme: solid line –set point; dashed line–expected open loop steady-state deviation

10.4.2 Experiments Using the SR Inferential Scheme

With the closed-loop experiments for the Brosilow scheme examined, attention was turned to the SR scheme. The below results employ the SR concentration predictor using 12 sensors and the 2 manipulated variables values and the maximal bed temperature predictor using parabolic interpolation of 3 measurements.

Stability Results. The stability of the closed loop system can be used to validate the uncertainty $\Delta_{c'm}$ (see Figure 10.2). Stability was tested for the conditions corresponding to experiments 4,5 and 6 in Table 10.3. The system was stable for operating conditions 5 and 6 but unstable for operating condition 4. This was due once again to the recycle flowrate becoming less than the minimal recycle flowrate selected for identification of the system. Therefore as concluded before, as long as the

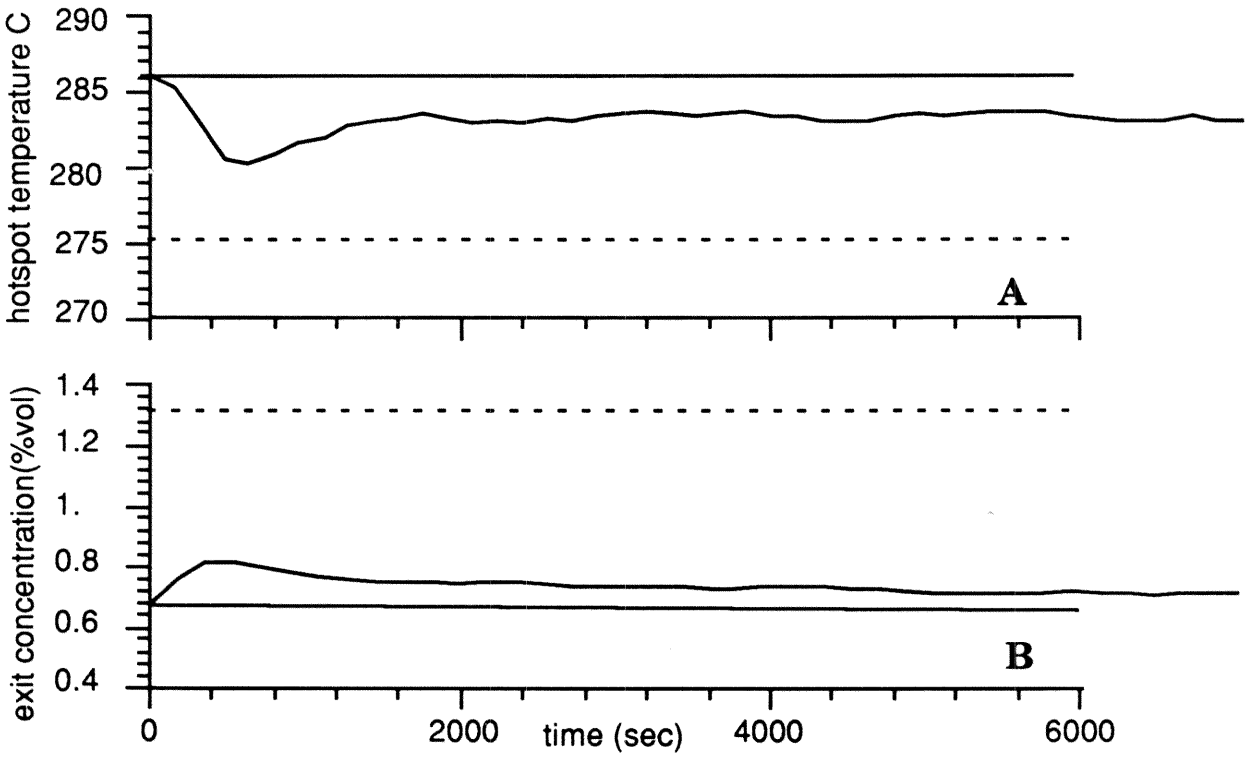

Figure 10.5. Response of controlled variables to step decrease in wall temperature (experiment 5), using SR inferential scheme: solid line – set point; dashed line – expected steady-state open loop deviation.

manipulated variables remain in the preselected window of operation for which the system was identified, the stability condition obtained with μ analysis was verified and consequently the uncertainty $\Delta_{c'm}$ was validated.

Performance Results. The time evolution of the controlled variables for experiment 5 in Table 10.3 is shown in Figure 10.5. A steady-state offset of 0.05% vol CO_2 was obtained for the concentration while an offset of approximately $3°C$ was obtained for the maximal temperature. The steady-state offset in concentration was four times less than the offset produced by the Brosilow predictor. This result is consistent with the open-loop results in which the SR predictor was more accurate at steady-state throughout the considered window of operation.

The predicted concentration using SR regression was compared for experiments

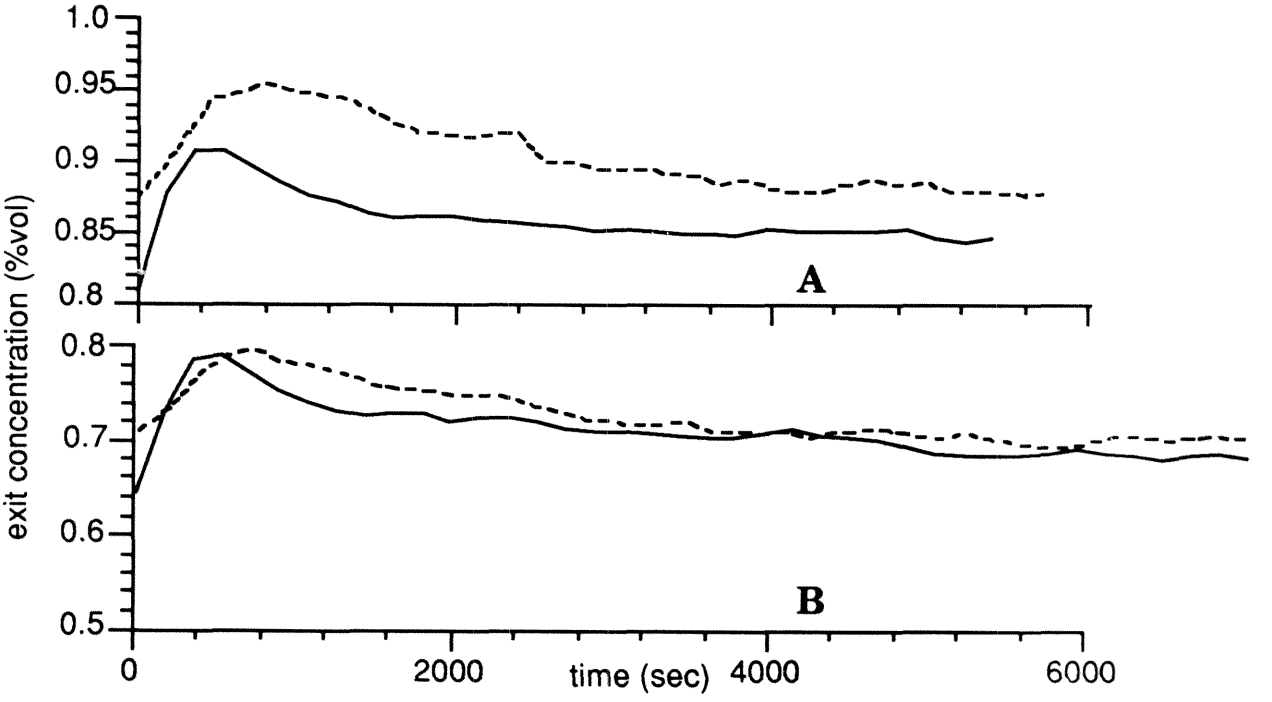

Figure 10.6. Comparison between exit concentration predicted by SR predictor (dotted lines) to concentration measured by GC (solid line) for: a- experiment 5, b- experiment 6.

5 and 6 to the actual concentration measured during these experiments. The results are shown in Figure 10.6. The maximal deviation obtained between the two variables was approximately 0.06% vol CO_2. This is quite accurate when one considers that the accuracy of the gas chromatographic system was approximately $\pm 0.05\%$ vol CO_2. These results partially validate the dynamic correction of the SR predictor using lead-lag compensation.

10.5 Conclusions

Two inferential control schemes were applied to an experimental methanation reactor. The first scheme relied on a Kalman filter-based approach, while the second used restriction regression. A SR predictor was shown to be significantly more accurate

for predicting the actual concentration in a wide range of operating conditions. A main disadvantage of the SR regression technique as used in this study was that only steady-state information was used in the regression process. Thus predictor had to be corrected to account for dynamic data via a simple lead-lag compensator. The zero and pole of this compensator were determined by fitting the predictor to experimental data.

The SR predictor was more accurate than Brosilow's predictor. With Brosilow's predictor the transfer functions G_{sm}, G_{sd}, G_{cm} and G_{cd} are separately identified and then the regression coefficients for the temperatures and the manipulated variables in the predictor are obtained from algebraic operations on these transfer functions. In the SR technique the coefficients are identified directly from the experimental data. Thus, the separate identification of the above mentioned transfer functions was bypassed and the relevant correlations are captured.

An additional difficulty with Brosilow's scheme arises from selecting the controller as the inverse of the process model. Due to the high condition number of the process under consideration, this type of control was sensitive to model uncertainty. In order to overcome this problem a decentralized controller was selected and its parameters tuned for robustness using the structured singular value. The use of this controller in conjunction with the SR predictor results in good closed loop performance. Moreover, with Brosilow's inferential scheme a larger number of measurements does not necessarily result in better prediction of the concentration for different operating points. However for the SR regression scheme the accuracy was shown to improve as the number of sensors was increased.

Developing a realistic but not overly conservative uncertainty description is crucial for analysis and tuning. In the SR scheme, a lumped uncertainty between the manipulated variables and the predicted controlled variables is identified directly. A design using the uncertainty associated with each one of the measurements separately would

be very conservative. Identifying correlations between the uncertainty elements can reduce the conservativeness of the design. This type of correlation was seen in the context of the Brosilow scheme between the elements of Δ_{sm}. On the other hand the SR scheme apparently accounted for all the correlations between the parameters and no correlations were found between the elements of the uncertainty $\Delta_{c'm}$.

Note

Versions of this chapter originally appeared in [7] and [8]. Hector Budman performed the experimental work and designed the Brosilow-type controller.

10.6 Appendix: Nomenclature Unique to Chapter 10

Variable	Description
c	Primary measurement/controlled variable
c'	Controlled variable normalized by performance weight W_p
c''	The product $c\ c_p^*$
d	Unmeasured disturbances
d'	Unmeasured disturbances normalized by performance weight W_d
G_{ij}	The transfer function from variable j to variable i
G'_{ij}	The transfer function at normal operating conditions
K_i	$i = 1, 2$. Proportional control gain in equation (10.13)
m	Manipulated variable
M	Interconnection matrix for structured singular value (SSV) analysis defined in Fig 4b
M_{ij}	The i,jth element of M
n	Noise to represent the offset between measured and predicted concentration
s	Secondary variables
T_s	Sampling time
W	Projection matrix of X
$W_{e,i}$	$i = 1, 2$. Weights used for SSV analysis.
W_d	Disturbance weight
W_p	Performance weight
δ_j	Deviation for the variable j
Δ_u	Uncertainty matrix
Δ_i^*	Individual uncertainty elements
λ	IMC filter constant
μ	Structured Singular Value (SSV)
τ_1	Controller integral reset time in equation (10.13)

Chapter 11

Metabolic Control Analysis

11.1 Introduction

The prior two chapters presented examples where significance regression provided an advantage over existing regression methods. This chapter will also present such a result, but more importantly will highlight limitations of current regression practice (including PLS) and demonstrate how the SR methodology provides a comprehensive framework for addressing these concerns. The experimental studies presented in this chapter (both "wet" and numerical) preceded the theoretical developments of Part III. Indeed, the topics in Part III were pursued in response to the difficulties encountered while conducting this study.

The problem studied was the modeling of the sensitivity of certain intra-cellular concentrations to small changes in various variables; in effect, one desires to compute the Jacobian of a living cell. Prior experimental work had been performed to build a comprehensive model of the relevant processes [17]. However, the work encountered severe numerical difficulties since the input variables could not be independently varied. Thus, the "wet" experiments were halted until the collinearity problem could be addressed. This problem was studied via simulation; data was generated from

and tested against solutions to the nonlinear system of kinetic equations described in appendix D. As shown in the third section restriction regression led to better results than were obtained with OLS. More important than the numerical outcome, however, was the insight the study yielded on the interaction between issues tractable with current theory and the needs of application. The specific problems encountered and the use of the significance regression framework to address these problems are discussed in the final section.

11.2 Metabolic Modeling

The motivation for this work has been concisely stated by Paul Schlosser, one of the co-investigators. Rather than dilute his biochemical expertise, his insight is quoted directly:

> In biochemical engineering, one is often interested in increasing the flux to, or steady-state concentration of, some cellular product. Although many enzymes may be involved in a metabolic pathway, experience indicates that there will be a relatively small number of enzymes (one, two, or maybe three) which "control" the flux or concentration of interest; i.e. increasing the concentration(s) of this small subset of enzymes will significantly increase the amount of product. Metabolic control analysis seeks to identify these enzymes by use of a model of the process under study which can be obtained from a knowledge of the pathway structure and kinetics, so as to avoid "trial and error" genetic engineering. [80]

The application of Metabolic Control Analysis (MCA) [16, 28, 48, 49, 74, 50] or other methods [78, 79, 81] of determining metabolic sensitivity to the rates of specific cellular processes, such as enzymatic reactions, requires knowledge of the elasticity coefficients (system partial derivatives) for the processes under study. Although rate

equations are available in the literature for some enzymatic reactions, there are many reactions and processes for which this is not the case. While one could perform the experiments necessary to determine the rate equations for a given system, these equations are, in fact, not required for the calculation of sensitivities – only the elasticities (the derivatives) are needed. A more direct and efficient approach would be to compute elasticities directly from experimental data. Errors can arise in such an approach, however, since data are invariably corrupted by noise and the data are typically collinear.

11.2.1 The General Metabolic Modeling Problem

The modeling of metabolic pathways or other biosynthetic processes, up to and including the development of structured models to describe over-all cellular growth and regulation, can be divided into two stages. In the first stage, the structure of the model is determined by listing a set of key metabolites or metabolic pools, and then identifying the reactions or "steps" which link those pools. In the second stage of modeling the rate or behavior of each of the "steps" is quantified. The behavior of a step might be quantified by assigning a rate equation or equilibrium relationship to that step. For the purpose of obtaining flux and concentration control coefficients [49], however, one only needs the elasticities (derivatives) of the rate through each step with respect to the set of metabolites or metabolic pools in the model, and these elasticities can be calculated directly from experimental data.

The determination of elasticities directly from data has been considered previously by Canela *et al.* [9] who discussed the possibility of experimentally manipulating the concentrations of individual metabolites *in vivo* by adding the metabolites to the growth medium. There are two difficulties with such an approach. First, many metabolites are not capable of being transported across the cell wall, so one could not effectively manipulate their concentrations in this manner. Secondly, even when

a metabolite is taken up by a cell, it's addition will probably cause the concentrations of other metabolites to vary at the same time because of cellular regulation (e.g., feed-back inhibition). Hence, one must consider the simultaneous changes of all precursors, cofactors and effectors for a given step when determining the elasticities from experimental data; thus one is faced with a collinear multi-variable regression problem.

Before discussing the evaluation of elasticities, a few definitions are necessary. Consider a subset of cellular metabolism that consists of a number of enzymatic reactions or other elementary cellular processes, along with those metabolites or pools that are the intermediates for these reactions. This subset could consist of a single enzymatic reaction with no intermediates. One could, given rate expressions for all of the reactions in this subset, write a detailed model describing its behavior. For a single enzymatic reaction, this "model" would just be its rate expression. The quantities, such as metabolite concentrations, that appear in this model can then be divided into dependent variables, such as the intermediate concentrations and reaction rates, and independent variables, such as the concentrations of precursors, cofactors and effectors that are not intermediates. The set of independent variables for a given metabolic subset can then be called its inputs and the set of dependent variables its outputs. These definitions follow from the fact that one can typically solve the model to obtain values for the dependent variables given values of the independent variables. An example of a metabolic subset is the "Cellular Metabolic Environment" depicted in Figure 11.1, with inputs and outputs indicated. The details of this model are described in Appendix D. Even if the model for some metabolic subset is unknown, one can still explore the relationship that exists between the inputs and outputs. This approach generalizes and complements the results obtained by Brown *et al.* [6] Let x and y be the vectors of inputs and outputs, respectively, of a metabolic subset, for which x_0 and y_0 are one pair of corresponding steady-state values. If the steady-state

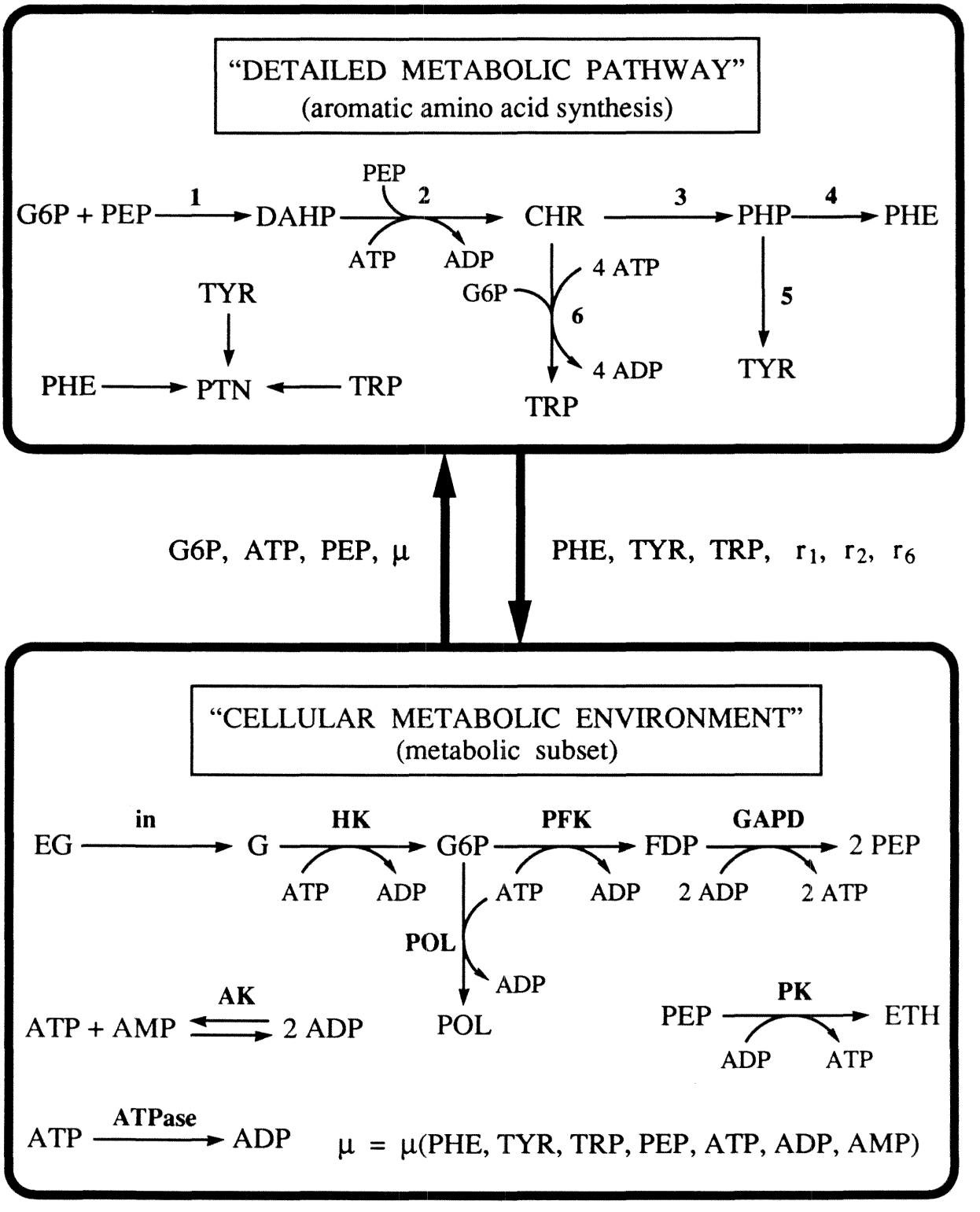

Figure 11.1. Metabolic processes included in a cell mass model developed by Schlosser and Bailey[81].

value of y is a unique function of x in the region of (x_0, y_0), then a Taylor series expansion about (x_0, y_0) yields

$$y - y_0 = (x - x_0)^T R + \mathcal{O}(\|x - x_0\|^2). \tag{11.1}$$

In biochemical engineering parlance, R is the "elasticity" of y to x for this metabolic subset; mathematically, R is the Jacobian of the mapping from x to y near x_0. If one has measured sets of corresponding values of x and y near the steady-state of interest, as well as x_0 and y_0, then equation 11.1 is equivalent to the classical regression problem, equation 2.1, and one can estimate R via linear regression. For a more detailed description, see Schlosser and Bailey [81]. Similarly, by first taking logarithms of the data, such an approach could be used to compute power-law coefficients for biochemical systems theory (BST) [78, 79]. Two difficulties with this approach are that one tends to be limited in the number and/or type of experiments that can be performed, and that the components of x are typically correlated. Thus, the classical least squares approach leads to unreliable estimates of the "elasticities."

11.2.2 The Specific Metabolic Modeling Problem

This study examines the steady-state behavior and sensitivities of the system in Figure 11.1 as modeled by in appendix D, and in particular examines the elasticities of the "Cellular Metabolic Environment." This metabolic subset provides precursors and cofactors for the "Detailed Metabolic Pathway" (the aromatic amino acid pathway). The inputs to this subset (elements of x) are the concentrations of the amino acids, tryptophan (TRP), tyrosine (TYR) and phenolalanine (PHE), along with the pathway fluxes r_1, r_2 and r_6. The outputs from this subset (elements of y) are the specific growth rate, μ, and the concentrations of glucose-6-phosphate (G6P), phosphoenol pyruvate (PEP) and adenosine tri-phosphate (ATP). The study proceeded

under the assumption that the only feasible observations of the "Cellular Metabolic Environment" were measurements of x and y. Two types of controllable "disturbances" were used to explore the relationship between x and y . First, step 1 in the aromatic amino acid pathway was inhibited to various extents. Second, the three amino acids (TRP, TYR or PHE) were "fed" to the system; one could also combine several of these disturbances in a single experiment. By measuring the steady-state values of x and y for various sets of disturbances, as well as the undisturbed values, one can form X and Y matrices as described above. The problem of determining a value for R for this model was first studied in Schlosser and Bailey [81]. The results described there, however, were obtained using "measurements" with six significant figures; such precision is not generally possible. In this study the following precision limitations are assumed: [TRP], [TYR] and [PHE] can be measured only to $\pm$0.5mM; r_1, r_2 and r_6 can be measured to $\pm$0.005mM/min; [G6P] and [ATP] can be measured to $\pm$0.005mM; PEP can be measured to $\pm$0.0005mM; and μ can be measured to $\pm 0.003 \text{hr}^{-1}$. These "inaccuracies" give each measurement two significant figures except for [TYR], [PHE] and the fluxes, which had three significant figures. The "training set" for the study was designed to be similar to the set of experiments described in Schlosser and Bailey [79]. The $\tilde{R}$'s computed by the various regression methods were tested using the calculations described in Appendix D.2 to obtain estimates of the concentration control coefficients (*ccc*'s) for the entire model; this approach assumes that the "detailed metabolic pathways" are known.

For the set of experiments described in Schlosser and Bailey [81, pp. 106 and 107, Table 1] and the precision described above, all of the regression techniques studied yielded erroneous *ccc*'s. A readily identified source of difficulty was that the signal-to-noise ratio (SNR, the measurement minus the nominal value, divided by the "measurement noise") of the inputs and the outputs are near unity. One is faced with a trade-off: signals should be as small to maintain linearity but large enough to

overcome the noise. Since (1) noise sensitivity is seen to be a severe problem, and (2) the problem is inherently collinear, restriction regressors such as PLS and SR are a promising approach for reliably estimating the elasticities and providing usable *ccc*'s.

11.3 Numerical Results

Since the existing experiments were too noisy, the simulation model was used to generate a new set of "experiments" with a better SNR; these are shown in Table 11.1. These perturbations were selected to be just large enough to give each measured variable a signal-to-noise ratio of at least 5 in one or more experiments, while staying approximately in a range of linear response. In general, the lowest SNR was for [PEP], so it was a limiting factor in the choice of perturbation sizes; the experiments gave a SNR of 4 for [PEP] in 3 experiments and a SNR of 6 in one other. Also, for simplicity, the percentage rate of addition of the amino acids was kept the same for all experiments in which addition occurred (40%), and the percent inhibition of r_1 was kept the same for all experiments in which step 1 was inhibited (20%). For "wet" experiments, some preliminary experiments are necessary to help establish "reasonable" perturbations. The X, Y, and the nominal values for these experiments, subject to the presumed errors listed above, are given in Table D.2 of appendix D. Variations in [G6P] were below the "detectable" level; hence [G6P] was taken to be constant and the corresponding row of R filled with zeros.

The exact values of the *ccc*'s were computed via numerical differentiation from the simulation model and are shown in Table 11.2. Likewise, the *ccc*'s estimated using ordinary least squares (OLS) are shown in Table 11.3. Although most of the *ccc*'s calculated by this method are in reasonable qualitative agreement with the actual values, the control of [TRP] (the r_6 column) fail to reflect the true importance of this step. Since a primary purpose of metabolic analysis is to identify steps which exert

Experiment No.	Feed Rates			Inhibition of r_1
	PHE	TYR	TRP	
1	40%	-	-	-
2	-	40%	-	-
3	-	-	40%	-
4	-	-	-	20%
5	40%	-	-	20%
6	-	40%	-	20%
7	-	-	40%	20%

Table 11.1. Experimental perturbations used in the MCA study.

ccc's	r_1^{max}	r_2^{max}	r_3^{max}	r_4^{max}	r_5^{max}	r_6^{max}
[TRP]	-0.166	-0.433	-0.209	0.037	-0.036	0.404
[TYR]	0.255	0.251	0.251	-0.512	0.513	-0.051
[PHE]	0.126	0.124	0.124	0.444	-0.433	-0.025

Table 11.2. "True" concentration control coefficients (*ccc*'s)

ccc's	r_1^{max}	r_2^{max}	r_3^{max}	r_4^{max}	r_5^{max}	r_6^{max}
[TRP]	-0.116	-0.097	0.020	0.071	-0.070	0.098
[TYR]	0.197	0.290	0.103	-0.517	0.518	0.147
[PHE]	0.097	0.143	0.051	0.441	-0.441	0.072

Table 11.3. Concentration control coefficients (*ccc*'s) computed using OLS.

ccc's	r_1^{max}	r_2^{max}	r_3^{max}	r_4^{max}	r_5^{max}	r_6^{max}
[TRP]	-0.135	-0.349	-0.450	0.031	-0.030	0.758
[TYR]	0.187	0.498	0.250	-0.492	0.493	-0.046
[PHE]	0.092	0.245	0.123	0.454	-0.454	-0.023

Table 11.4. Concentration control coefficients (*ccc*'s) computed using PLS.

the greatest level of control on a metabolic parameter, OLS is clearly inadequate.

Next studied was the PLS restriction regressor reviewed in appendix A. One should notice there there are six inputs to the model ($n_i = 6$), but only four degrees of freedom for conducting the experiment. Thus Rank(X) can exceed four only due to noise or nonlinearity; clearly, restriction regression is appropriate. Interestingly, a variety of restriction regression methods, including principal components analysis, PLS, and a clever *ad hoc* method that had been developed earlier all lead to the almost identical W with Rank(W) = 4. The details of how these methods were implemented will not be labored here because all of those techniques (except cross-validation on the PRESS) have since been superseded by the results of Part II; suffice here to say that the data spoke very strongly in favor of a four-dimensional description, regardless of which restriction regressor was used and what method was used to determine n_d.

The *ccc*'s computed using PLS are shown in Table 11.4. The improvement in the trace of the MSE was not dramatic: 0.59 for OLS versus 0.52 for PLS. However, PLS yielded more realistic *ccc*'s for the purposes of metabolic control analysis; in particular, PLS properly indicated the strong influence of r_6 on [TRP].

11.4 Discussion of Key Issues

As shown in the previous section, use of the PLS restriction regressor leads to a more effective estimation of the concentration control coefficients. However that is actually one of the less interesting interesting outcomes of the study. Of greater importance is the light this study shed on the then-state-of-the-art PLS algorithm and the general problems of implementing restriction regressors in the "real world." One of the most pressing questions in this study was understanding the impact of the measurement errors and how to craft the analysis to minimize their impact. Addressing this question required the study of estimation intervals, measurement

error models, and scaling. After these questions are addressed, another slate of issues arose associated with the question "how do I design the next experiment." What few results were available at the time of the study [62] were generally *ad hoc* in nature and did not use a consistent framework; with the significance regression method, some of these questions can be answered outright, and all of them can be approached and discussed within a consistent framework.

Effect of input errors. When this study was undertaken the effect of the input measurement noise was largely unknown; in fact, some of the early work labored under the assumption that these errors, not the collinearities, were the dominant source of inaccuracies in the estimated *ccc*'s [81]. The effect of these measurement errors was analyzed in chapter 6 and will not be repeated here, except to state that

$$\mathrm{Var}(\tilde{b}) \approx \sigma_e^2 W(W^T X^T X W)^{-1} W^T + W(W^T X^T X W)^{-1} W r r^T W^T (W^T X^T X W)^{-1} W^T \tag{11.2}$$

when significance scaling is used. Regardless of input errors, the variance will be "large" whenever $(W^T X^T X W)$ exhibits collinearities and will be proportional to σ_e^2 in all but one direction. Only in the one direction, $W(W^T X^T X W)^{-1} W r$, did the input errors have any effect on the variance. Thus one can now see reducing the input measurement errors is beneficial, but reducing σ_e^2 will typically have a much greater impact and managing collinearities is essential.

Error bounds. Few experimentally obtained numbers are useful unless one can also state the error bound. The error bounds can be obtained from the interval estimates discussed in chapter 2. Alas, SR does not improve the error bounds over the classical error bounds unless one can assure that $r \in \mathrm{Range}(W)$. Since the experimental protocol limits the degrees of freedom to less than n_i, one cannot demonstrate unbiasedness in the example studied here. Still, this vital issue is readily examined within the SR framework.

Scaling. The strong influence scaling has on restriction regressors has long been known and was discussed in chapter 8. For this problem, the experimenters explicitly wanted to know if scaling could be used to mitigate the effect of the input measurement errors. The literature contained few suggestions [62] and no clear answers. While the scaling issue has not been conclusively solved, the results of chapter 6,the discussion of sec 8.6, and the results of section 9.5 provide support for the use of significance scaling with SR to reduce the effects of measurement noise.

Choice of the next experiment. If the current *ccc*'s are determined to be inadequate (say, due to the error bounds being too large), how does one choose the next set of experiments? How many experiments are useful? If errors can be reduced, which ones should be? This last point is particularly crucial. Often, various measurements can be made more precise by using more elaborate laboratory procedures, repeated experiments, or stronger radio-active labels. The first two involve time and expense, and the third increases health risks. Thus, a strong analytical framework to help design the "next" experiment is vital. The issue of experimental design is a topic unto itself and will not be properly discussed here; only the few additional insights into metabolic modeling provided by this study are presented here. The interested reader is referred to standard texts such [3, 4] for a more comprehensive discussion.

The most important contribution from significance regression has already been discussed: reducing the measurement errors in the outputs is generally more important than reducing the measurement errors in the inputs. Moreover, one can compute the effect of additional experiments on the errors bounds; this allows one to easily compare the benefits of different contemplated experiments. The procedure is straightforward. One would calculate the x that results from a contemplated perturbation, build the matrix $X^T_{new} = [X^T|x]$ and compute the various error bounds discussed in chapter 2.

For a given noise level, one would generally choose the perturbations of additional

experiments so as to make the X as "well-rounded" as possible. Let n_f be the number of perturbations available (and therefore the degrees of freedom of the problem). If $n_f = n_d$, then one should choose the perturbations such that the x produced is proportional to eigenvector of minimum eigenvalue of $W^T X^T X W$. This is the currently least informative direction being used for regression, so information in this direction will have the greatest impact on the error bounds. If $n_f > n_d$, one needs to establish if this problem is due to the mapping from the perturbations to the x being singular or due to excessive noise. The singularity of the (linear) mapping can be investigated by performing orthogonal design on the perturbations (as was done in this study) and performing principal components analysis [45] on the resulting X. Since all perturbation directions were investigated, any principal component (eigenvalue of $X^T X$) that fails to exceed the noise level for that direction is clearly associated with an (output) singular direction in the mapping. If significance scaling is used, this condition can be tested easily. The easiest test is that the principal component is "significant" if $\lambda_i(X^T X) > n_s$ where λ_i is the ith eigenvalue. A more statistically sound test can be derived from the fact that $\lambda_i(X^T X)$ is associated with a singular direction (principal component). If the measurement errors are assumed to be normal, then $\lambda_i(X^T X)$ has a χ^2-distribution with n_s degrees of freedom when the corresponding "true" eigenvalue is zero. Yet more sophisticated tests are available [32]. In the singular case, n_f should be reduced to equal the degrees of freedom one actually has to manipulate the x's. If the mapping is non-singular and $n_d < n_f$, one should investigate the "next-most significant direction" by selecting perturbations that produces an x proportional to $w^{opt}_{n_d+1}$.

Heteroscedastic errors. The previous paragraph invoked homoscedasticity; that is, that the noise levels were the same for each experiment. Since the experimenter has some control over the measurement errors in any given experiment, this is a vital issue. Heteroscedastic output measurement errors are easily handled by out-

put scaling, as discussed in section 8.5. The theory developed in Part II and III does not rigorously hold for heteroscedastic input measurement errors, but, as discussed in section 8.5, the theory provides a useful guide. One of the chief benefits of the study was highlighting the need for better methods to treat problems with heteroscedastic input measurement errors. This problem is neglected in much of the literature due to it's analytical difficulty [102].

Evaluating the linearity assumption. The entire development has assumed linear relationships; yet, the relationships between the various metabolic quantities are known to be non-linear. Determining whether or not the experiments remained within the "approximately linear" range previously relied almost entirely on the experimenter's intuition. While such intuition is essential and should not be discounted, simple numerical tests are also useful. The analysis framework developed in this discussion provides a simple numerical test: if λ_{n_f+1} is significant, then the system is exhibiting non-linear behavior. This test is based on the observation that if the system is behaving linearly, then λ_{n_f+1} must consist solely of noise. However, as illustrated in the following brief example, nonlinearity can lead to more "significant" principal components than degrees of freedom. Assume that only one available perturbation is experimentally possible; for each experiment, this perturbation has value π_j. Moreover, the inputs respond to this perturbations as $x_j^T = [\pi_j \ \pi_j^2]$. Notice that all of the possible x_j lie on a parabola, so there is still only one degree of freedom. However,

$$X^T X = \begin{bmatrix} \sum_{j=1}^{n_s} \pi_j^2 & \sum_{j=1}^{n_s} \pi_j^3 \\ \sum_{j=1}^{n_s} \pi_j^3 & \sum_{j=1}^{n_s} \pi_j^4 \end{bmatrix} \tag{11.3}$$

which is generally not singular. The nonlinear response of the inputs to the perturbations made both eigenvalues of $X^T X$ positive despite the problem of only possessing one degree of freedom.

11.5 Conclusions

This chapter examined the role of regression in metabolic modeling. The general problem of metabolic control analysis was examined and a particular metabolic problem presented. The performance of OLS and PLS were studied via simulation; PLS was found to provide more physically useful concentration control co-efficients. More importantly, this study revealed many of the "real world" benefits of using the significance regression framework. These benefits included understanding the relative importance of input and output measurement errors, motivating particular scaling regimes, and developing simple numerical tools to evaluate nonlinearites and singularities in the mapping between the perturbations and the inputs.

Note

Versions of this chapter originally appeared in [82].

Part V

Summary, Recommendations, and Future Work

Chapter 12

Summary and Future Work

12.1 Summary of Contributions

This worked examined restriction regressors for collinear problems. These regressors had the form

$$\tilde{b} = \arg \min_{v \in \text{Range}(W)} (y - Xv)^T (y - Xv) \tag{12.1}$$

where $W \in \Re^{n_i \times n_w}$. The general prediction and estimation properties of this class of regressors were systematically studied and the criterion for a restriction regressor to be "good" was elucidated. This concept of "good" translated directly into a null hypothesis. A test statistic was derived for this hypothesis, and a novel regression procedure, significance regression (SR), was derived by maximizing this test statistic. The successful but incompletely understood PLS algorithm was shown to a computational method for performing SR for scalar output problems. The statistical basis for SR made possible the determination of the dimension of the search space (the "number of directions") via significance testing; prior PLS approaches had used cross-validation on the PRESS exclusively. While SR was derived for scalar output regression problems, SR can also be described in general terms using linear operator theory; in this general form, SR also gives rise to a factor analysis method. The linear

operator description allows the basic features that make SR successful to be directly transferred to other problem domains.

The SR method was extended to vector output problems. The PLS algorithm was shown to be closely related to the SR factor analysis algorithm but sub-optimal for regression. SR can be just as easily extended to tensor problems via the linear operator description. The SR method was applied to the measurement error model, resulting in a novel estimator for this important class of problems. Further analysis of collinear problems showed that methods based on the "classical" model, which does not take into account input measurement errors, can also be successful. The regressors developed thus far were found to have poor robustness properties, and thus were sensitive to outliers. A novel weighted least-squares significance regression method was derived to take advantage of the robustness properties of the M-estimators. A novel robust SR algorithm resulted that maintained SR's strengths for handling collinearity while inheriting the robustness properties of the M-estimators. These theoretical developments were then used to address the difficult but vital issue of scaling. Significance scaling was seen to be a useful adjunct to significance regression, and the importance of rendering the data homoscedastic was reinforced.

A suite of simulations was performed that verified various points made in the course of the theoretical analysis. Vector output problems, measurement error model problems, problems with outliers, and scaling methods for these problems were each examined in turn. Restriction regression was used in the construction of a robust inferential controller for a packed-bed reactor. The resulting control scheme employed a better predictor than previous literature recommendations provided, used an inferential controller that embodied fewer linearity assumptions than a Kalman filtering-based approach, and produced better performance than previous approaches. Restriction regression was also applied to the modeling of cellular metabolism. Here the objective was to compute the Jacobian of certain key cellular processes. Restric-

tion regression provided the means to forge ahead where the previously attempted methods had failed. Study of the analysis, simulations, and applications described in this thesis led to definite recommendations for regression practitioners and to several promising avenues of further inquiry.

12.2 Recommendations

Many of these recommendations are consistent with current practice; none of the suggestions will seem counter-intuitive to most practitioners. The recommendations made here are not on equal footing. For example, the robust regression methods endorsed here stand on firm theoretical ground supported by extensive "real world" use, while the scaling suggestions should be taken as just that: suggestions. Each step refers to a particular section; the reader is encouraged to review the referenced section for further explanation of the recommendation, discussion of the reasons supporting the recommendation, and the caveats that accompany the recommendation.

Algorithm 12.1 (Recommended significance regression procedure)

1. If one has strong *a priori* beliefs about the "true" distribution of r, do not use this algorithm. Instead use the appropriate Bayesian method. (Section 2.1)

2. If the problem has multiple outputs, "stack" data into a scalar output problem. (Section 5.2)

3. "Center" the data using nominal values if such values are meaningful and available, or the sample mean if a nominal point is not available. (Section 2.1)

4. Output scale the data to make the output errors (e) homoscedastic. The output scaling is D_2 of equations 8.1 and 8.2. The computation of D_2 is described in equation 8.10. (Section 8.5)

5. Input scale the data using significance scaling. The input scaling is D_1 of equation 8.1. If the inputs are not subject to measurement errors, or if the covariance structure of the input errors is completely unknown, use auto-scaling. (Sections 8.4, 9.3, and 9.5).

6. Select a robust M-estimation objective function. Use Huber's "proposition 2" objective function unless the errors are known to be "heavy tailed." In this case use Tukey's bi-weight objective function. (Section 7.3)

7. Choose a method for computing n_d. For estimation problems, the significance tests of section 3.2 can be effective and are computationally inexpensive. For prediction, cross-validation on the PRESS is the clear choice. (Sections 3.2 and Chapter 9)

8. Compute the restriction regressor, $\tilde{b}$, via robust significance regression, algorithm 7.3. If one is using cross-validation on the PRESS, "classical" significance regression, algorithm 3.1, can also be effective. (Sections 7.4 and 9.4)

9. Evaluate the resulting regressor. Do not accept $\tilde{b}$ if $\tilde{b}$ is "unreasonable," for no statistical method is infallible.

12.3 Suggestions for Future Work

The results of this inquiry revealed vital areas that call for further work and illuminated several promising avenues for future research; as so often happens in research, these areas only partially overlap. Chapter 2 clearly calls for further development. Now that the three most popular restriction regressors are readily described within a simple framework, what more can be learned? In particular, stepwise regression is widely used for variable selection, despite known deficiencies (the procedure is only

locally optimal). One should be able to use the SR viewpoint to produce an "improved" variable selection method. Equally intriguing will be clearly delineating the relative strengths of SR and PCR. This is still an open question, as illuminated by recent comments of Næs [67] and Joliffe [46]. A point of considerable interest to statisticians was unearthed but not pursued in this study. In section 2.3, the analysis revealed that the component of r in the "rejected" subspace still produced a bias *in the search space*. The wider implication of this finding is that even the much-vaunted "unbiased" estimator is biased if *any* input that affects the output and correlates with any included input is omitted from the model. When one also considers the deficiencies of the unbiased estimator for collinear problems and for the measurement error model, one might begin to question if "unbiasedness" is actually a useful criterion for regressors, or simply an easily evaluated criterion.

In breaking new ground for methods to determine n_d, this work revealed much work to be done. The current significance test is based upon a flawed independence assumption; for "clean" problems where n_d is "large" the cumulative errors due to this bad assumption might be fatal. Also, more appropriate tests need to be developed for prediction objectives. The other vital area is scaling, particularly for heteroscedastic input measurement errors. This study did find a new theoretical suggestion in favor of significance scaling, but much work remains. The solution path probably leads to Bayesian analysis, and in particular to understanding robustness to incorrect priors. Encouragingly, strong results already exist in this area. For example, chapter 8 of Gruber [24] reviews these results.

In the past few years there have been several notable *ad hoc* attempts to extend PLS to polynomial and other nonlinear models [73, 100]; however these methods all have major deficiencies [36]. The results of Chapter 4 point to several new avenues of attack. The simplest is to form additional columns of X from various monomial functions of the inputs and use the existing SR method. The historical problem with

such an approach has been the collinearity of the resulting X and the admission of too many degrees of freedom to the fitting procedure. SR mollifies the collinearity problem; moreover, SR admits only n_d degrees of freedom (roughly) to the fit, potentially also mollifying polynomial's notorious propensity for "over-fitting." The brief Fourier illustration at the end of chapter 4 points to another promising approach: combining algorithm 4.1 with various classes of basis functions such as Tchebyshev polynomials or radial basis functions [72] to produce novel interpolaters. Since SR explicitly addresses collinearity problems, the basis functions need not be orthogonal.

Shifting focus from theoretical considerations to applications, two key areas seem ripe for progress. As discussed in Chapter 11, much remains to be done on the metabolic modeling problem. The first order of business is to apply the results of this thesis to the current difficulties. If the remaining modeling obstacles can be cleared (a substantial if), considerable intellectual and fiduciary profits await. The other application calling for attention is model predictive control (MPC). One of MPC's chief advantages is that the controller is based on a model regressed directly from the data; the usual wall between "identification" and "control" that pervades most theoretical control analysis is relatively low for MPC. Historically, MPC has used Bayesian point predictors, although Ricker [75] has shown that restriction regressors are also adequate. Using the prediction interval results of section 2.3 one should be able to incorporate more information from the process data into the controller, thereby increasing reliability and robustness for these methods. With such an approach, the "nominal model" and the "uncertainty" are computed together, potentially providing a new "robust modeling" tool. Moreover, the tools of econometrics [12, 93] can be be applied to the regression process to include the effects of saturation; current modeling methodologies typically neglect this source of bias. Lastly, improvements in the significance tests should improve the process for choosing n_d for MPC models. Currently the PRESS is used on open loop data to determine n_d. One may be able

to formulate "control relevant" significance tests that better reflect the closed-loop inputs the model will face in operation.

Appendix

Appendix A

PLS Review

This section restates known results for partial least squares in a manner more compatible with the results developed in this paper. The algorithm was primarily developed by Wold [101]. A tutorial was provided by Geladi and Kowalski, [20] while Höskuldsson [39] analyzed the mathematical aspects of the algorithm. Helland [29] illustrated important properties of the loading vectors (w_i's). The view of PLS developed below draws strongly from Höskuldsson's and Helland's work.

As described by Höskuldsson, the PLS vectors at a given step of the process are found by computing the eigenvector of maximum eigenvalue in the following equations:

$$X^T Y Y^T X w^{opt} = \lambda_w w^{opt} \tag{A.1}$$

$$Y^T X X^T Y c^{opt} = \lambda_c c^{opt}. \tag{A.2}$$

Another way to describe w^{opt} and c^{opt} is

$$\{w^{opt}, c^{opt}\} = \arg \max_{||w||=1,\ ||c||=1} < Xw, Yc > \tag{A.3}$$

which is equivalent to

$$\{w^{opt}, c^{opt}\} = \arg\max_{w,c} \frac{w^T X^T Y c}{\sqrt{w^T w c^T c}}. \tag{A.4}$$

Taking the gradient of equation A.4 with respect to w and equating the result to zero yields

$$w^{opt} = \gamma X^T Y c^{opt} = \gamma X^T X \tilde{R} c^{opt}, \tag{A.5}$$

where γ is an irrelevant (but known) scalar. Thus, to compute the first PLS vectors, one solves the eigenvector equation A.2 to determine c^{opt} and then uses A.5 to produce w^{opt}. For the case of scalar output, c is a scalar so $w^{opt} = X^T y = X^T X \tilde{r}$.

Once the first PLS vector has been determined, w^{opt} is orthgononally removed from the input data, and the effect of w^{opt} and c^{opt} is subtracted from the output. The algorithm can thus be described:

Algorithm A.1 (Generation of PLS loading vectors)

$X_1 = X$ (A.6)

$Y_1 = Y$ (A.7)

Do $i = 1$ to n_d

Let c_i^{opt} be the eigenvector of maximum eigenvalue in

$$\lambda_{c_i} c_i = Y_i^T X_i^T X_i Y_i c_i \tag{A.8}$$

$$c_i^{opt} = \frac{c_i^{opt}}{||c_i^{opt}||} \tag{A.9}$$

$$w_i^{opt} = \frac{X_i^T Y_i c_i^{opt}}{||X_i^T Y_i c_i^{opt}||} \tag{A.10}$$

$$\beta_i = \frac{w_i^{opt\,T} X_i^T Y_i c_i^{opt}}{w_i^{opt\,T} X_i^T X_i w_i^{opt}} \tag{A.11}$$

$$\tilde{Y}_i = \beta_i X_i w_i^{opt} c_i^{opt\,T} \tag{A.12}$$

$$X_{i+1} = X_i (I - \frac{w_i^{opt} w_i^{opt\,T} X_i^T X_i}{w_i^{opt\,T} X_i^T X_i w_i^{opt}}) \tag{A.13}$$

$$Y_{i+1} = Y_i - \tilde{Y}_i \tag{A.14}$$

End Do

In PLS, one wishes to restrict the regressor to the input space and output space described by the PLS "loading vectors," w_i^{opt} and c_i^{opt}. Let $W = [w_1^{opt}|\ldots|w_{n_d}^{opt}]$ and let the orthonormal columns of C be such that $\text{Range}(C) = \text{Span}(\{c_1^{opt}, c_2^{opt}, \ldots, c_{n_d}^{opt}\})$. Moreover, define $C^{\perp}$ such that $[C|C^{\perp}]^T\,[C|C^{\perp}] = I$ and $n_c = \text{Rank}(C)$. The parameterization of all such allowable restriction regressors is $\tilde{B} = WVC^T$ where $V \in \Re^{n_d \times n_c}$. One solves for the least squared-error regressor in the classical manner.

$$
\begin{aligned}
V^{opt} &= \arg \min_{V \in \Re^{n_d \times n_c}} \|Y - XWVC^T\|_F^2 && \text{(A.15)}\\
&= \arg \min_{V \in \Re^{n_w \times n_c}} \| \, Y[C \mid C^{\perp}] - XWVC^T[C \mid C^{\perp}] \, \|_F^2 && \text{(A.16)}\\
&= \arg \min_{V \in \Re^{n_w \times n_c}} \| \, [YC \mid YC^{\perp}] - [XWV \mid 0] \, \|_F^2 && \text{(A.17)}\\
&= (W^T X^T X W)^{-1} W^T X^T Y C, && \text{(A.18)}
\end{aligned}
$$

from which the PLS restriction regressor is $\tilde{B} = W(W^T X^T X W)^{-1} W^T X^T Y C C^T$.

Appendix B

Routines To Generate Synthetic Data

All simulations were conducted using Matlab [64]. All simulations where started from the same random number seed (the default seed at startup), so simulations using the same code to generate problems used the identical one-thousand examples.

B.1 Vector Output Examples

The vector output examples of section 9.2 were generated using the `gendat2` routine in section B.5.1 with the parameters: `n_train = 30`, `n_test = 100`, `d = 10`, `o = 4`, `d_ind = 3`, `max_exp = 5`, `min_exp = 0`, and `noise = 0.5`.

B.2 MEM Examples

The MEM examples of section 9.3 were generated using the `gendat2` routine and code below. The parameters used were: `n_train = 30`, `n_test = 100`, `d = 10`, `o = 4`, `d_ind = 3`, `max_exp = 5`, `min_exp = 0`, `x_noise_max = 1`, `x_noise_min = 0.1`, and `o_noise = 0.2`.

```
% use data such that null pred has rms 1
    [X,y,Xt,yt,b] = gen_dat2(n_train,n_test,d,1,d_ind,max_exp,min_exp,o_noise);


% corrupt input data with noise
```

```
    rand('normal');



   W = diag(scaled_rand(x_noise_max,x_noise_min,d));


   rand('normal');
    X       = X +   rand(n_train,d)*W;
    Xt      = Xt+   rand(n_test,d)*W;
    W       = W*W;
```

B.3 Vector Output Examples with Outliers

The vector output examples of section 9.4 were generated using the `gen_dat2_rob` routine in section B.5.2 with the parameters: `n_train = 30`, `n_test = 100`, `d = 10`, `o = 4`, `d_ind = 3`, `max_exp = 5`, `min_exp = 0`, `noise = 0.5`, and `e = 0.5`, `outvar = 3`.

B.4 Scaling Examples

The MEM examples of section 9.5 were generated using the `gendat2` routine and code below. The parameters used were: `n_train = 30`, `n_test = 100`, `d = 10`, `o = 4`, `d_ind = 3`, `max_exp = 5`, `min_exp = 0`, `x_noise_max = 1`, `x_noise_min = 0.1`, and `o_noise = 0.2`.

```
P_nom = [ 3*ones(10,1); ones(10,1); ones(10,1)/3];

    rand('normal');

    X       = X +  diag(P_nom)* rand(n_train,d)*W;
    Xt      = Xt+   rand(n_test,d)*W;
    W       = W*W;

%  Compute the least squares regressor
r_temp = X\y;
```

```
%  compute the output scaling matrix
P_scale = diag( [
                ones(10,1)/sqrt(o_noise^2+   9*r_temp'*W*r_temp);
                 ones(10,1)/sqrt(o_noise^2+    r_temp'*W*r_temp);
                 ones(10,1)/sqrt(o_noise^2+    r_temp'*W*r_temp/9)]);
```

B.5 The Random Example Generation Routines

B.5.1 Routine to Generate Random Regression Problems

```
function [X,y,Xt,yt,b] =gen_dat2(n_train,n_test,d,o,d_ind
                                    ,max_exp,min_exp,noise)

% this function generates data for linear regression problems
%
%
%  n_train  is the number of samples to be the training set
%  n_test   is the number of samples to be the testing set
%  d        is the number of inputs
%  o        is the number of outputs
%  d_ind    is the number of inputs NOT rotated
%            and thus "independent"
%  max_exp  the largest order of magnitude contemplated
%  min_exp  the smallest order of magnitude contemplated
%            used for scaling the input data and
%            generating the regression vector
%  noise    std deviation of the normal additive noise
%
%
%  X         is the input training data
%  Xt        is the input testing data
%  y         is the output (noise corrupted) training data
%  yt        is the output (not noise corrupted) testing data


scale = diag(abs(scaled_rand(max_exp,min_exp,d)));
% these b's are for the same direction as singular vectors
for i=1:o
    b(:,i) = scaled_rand(max_exp,min_exp,d);
```

```
end

% need to build random orthogonal matrix
% only rotate d - d_ind columns; let the rest be
% 'approx' independent

d_rot =  d - d_ind;
if  d_ind == d
  v = eye(d);
else
  rand('uniform')
  v = rand (d_rot);
  [u,s,v] = svd(v);
  if d_rot == d
      v = u*v;
  else
      v = [ eye(d_ind), zeros(d_ind,d_rot); zeros(d_rot,d_ind), u*v];
  end

end

% use v as an additional rotation on the data and regression vector

rand('normal')
X =  rand(n_train,d) * scale * v;
Xt =  rand(n_test,d) * scale * v;
b = v'*b;
yt = Xt*b;
%desire RMS of null predictor to be 1
rms = sqrt(trace(yt'*yt)/ (n_test * o) ) ;
b=b/rms;
yt = Xt*b;
y = X*b + rand(n_train,o)*noise;
```

B.5.2 Routine to Generate Random Regression Problems with Outliers

```
function [X,y,Xt,yt,b] =gen_dat2_rob(n_train,n_test,d,o,d_ind,max_exp,min_exp,n

% this function generates data for linear regression problems
%
%
%  n_train  is the number of samplesto be the training set
```

```
%  n_test   is the number of samplesto be the testing set
%  d        is the number of inputs
%  o        is the number of outputs
%  d_ind    is the number of inputs NOT rotated and thus "independent"
%  max_exp  the largest order of magnitud contemplated
%  min_exp  the smallest order of magnitud contemplated
%            used for scaling the input data and
%            generating the regression vector
%  noise    std deviation of the nomral additive noise
%
%
%  X         is the input training data
%  Xt        is the input testing data
%  y         is the output (noise corrupted) training data
%  yt         is the output (not noise corrupted) testing data



scale = diag(abs(scaled_rand(max_exp,min_exp,d)));
% these b's are for the same direction as singular vectors
for i=1:o
    b(:,i) = scaled_rand(max_exp,min_exp,d);

end

% need to build random orthogonal matrix
% only rotate d - d_ind columns; let the rest be
% 'approx' independent


d_rot =  d - d_ind;
if  d_ind == d
  v = eye(d);
else
  rand('uniform')
  v = rand (d_rot);
  [u,s,v] = svd(v);
  if d_rot == d
      v = u*v;
  else
      v = [ eye(d_ind), zeros(d_ind,d_rot); zeros(d_rot,d_ind), u*v];
  end

end
```

```
% use v as an additional rotation on the data and regression vector

rand('normal')
X =  rand(n_train,d) * scale * v;
Xt =  rand(n_test,d) * scale * v;
b = v'*b;
yt = Xt*b;
%desire RMS of null predictor to be 10
rms = sqrt(trace(yt'*yt)/(n_test* o) );
b=b/rms;
yt = Xt*b;
y = X*b;
% need to produce outliers
for i= 1:n_train
   for j = 1:o
      rand('uniform')
      if (e < rand)
        rand('normal')
        y(i,j) = y(i,j) + rand*noise;
      else
        rand('normal')
        y(i,j) = y(i,j) + rand*noise*out_var;
      end
   end
end
```

B.5.3 Routine to Generate "Exponential" Random Numbers

```
function vect =  scaled_rand(u,l,d)

% this function generates a vector of random numbers that are
% 'exponentially' distributed;  that is, the probability of
%  a number having any given  order of magnitude within
%  the valid range is roughly equal
%
%   u     lowest order of magnitude allowed
%   l     highest order of magnitude allowed
%   d     is the dimension of the vector generated
%
%  10^l < number < 10^u
%
```

```
rand('uniform');

for i = 1:d
  vect(i) = 10^ ( (u - l) * rand(1,1)  + l);
end

vect = vect';
```

Appendix C

Raw Data from Packed Bed Reactor Experiments

This appendix contains the raw data used to build the predictors in chapter 10. The manipulated variables (recycle rate and heating power), the disturbance (wall temperature), and the controlled variable (CO_2 outlet concentration) are reported in Table C.1. The corresponding axial bed temperatures are in Table C.2. Here the temperatures are numbered T_1 through T_{12}. However, there were actually 15 thermocouples on the reactor, of which only the twelve reported here produced reliable readings. Thus "thermocouple 13," which was the single sensor used in the Brosilow scheme, corresponds to T_{10} in Table C.2.

Experiment No.	Manipulated Variables r slpm	h % of 1800 W capacity	Disturbance T_w $°C$	Output concentration %vol
1	17	5	257	0.74
2	18	50	254	0.65
3	13	45	256	0.18
4	18	20	255	0.99
5	19	10	255	1.3
6	19	10	256	1.18
7	20	35	255.5	0.93
8	17	50	257	0.24
9	19	85	255	0.17
10	15	45	256	0.26
11	19	10	254	1.41
12	19	10	257	1.06
13	19	35	256	0.81
14	16	35	255.5	0.53

Table C.1. Raw data for the 14 experiments used to build the PLS predictor for the methanation reactor.

Experiment No.	T_1	T_2	T_3	T_4	T_5	T_6	T_7	T_8	T_9	T_{10}	T_{11}	T_{12}
1	222	235	239	244	250	257	265	268	272	277	282	285
2	246	253	255	258	263	267	273	274	275	277	277	279
3	256	265	267	272	278	285	292	292	292	290	284	284
4	232	242	245	249	254	259	266	268	270	274	276	279
5	216	229	233	238	244	250	257	260	264	268	273	277
6	217	230	234	240	245	251	259	262	265	270	276	279
7	233	243	245	249	254	259	265	267	270	273	276	279
8	252	260	263	266	271	277	283	284	285	285	283	284
9	267	272	272	275	278	282	285	284	283	281	277	277
10	253	261	263	267	272	278	284	285	286	286	283	283
11	215	228	232	237	243	249	255	258	262	266	271	274
12	218	231	235	240	246	253	260	263	267	272	277	281
13	234	243	246	250	255	261	267	269	272	275	278	281
14	243	252	254	258	263	269	276	278	280	282	283	285

Table C.2. Raw temperature data ($^{\circ}C$) for the 14 experiments used to build the PLS predictor for the methanation reactor

Appendix D

Metabolic Model for MCA Study

The appendix discusses the model used in chapter 11 for simulating cellular metabolism developed by Schlosser, Galazzo and Bailey [17, 18, 81]. This appendix is adapted directly from Schlosser [82]. The first section presents the main model. The second section presents the "Black Frame" by which the perturbations are related to the inputs. The final section presents the numerical data used with the models. The nomenclature in the appendix deviates from the rest of the thesis so as to be as similar as possible to the nomenclature used in the biochemical engineering literature.

D.1 Cellular Metabolism Model

A set of reactions that represents the metabolism of a cell mass is depicted in Figure 11.1. The detailed pathway represents the immediate processes by which the aromatic amino acids are produced and utilized, which would be described by a detailed model of the pathway's behavior (amino acid transport between cells and growth medium not indicated). The cellular metabolic environment represents the set of all other cellular processes which impact the detailed metabolic pathway. The interaction variables are the quantities by which these two sets of processes communicate. The chemical species in chemical species Figure 11.1 are: G = extracellular glucose, G6P = glucose-6-phosphate, FDP = fructose-1,6-diphosphate, PEP = phosphoenol pyruvate, POL = polysaccharides, ETH = ethanol, DAHP = 3-deoxy-D-arabino- heptulosonate-7-

phosphate, CHR = chorismate, TRP = tryptophan, PHP = prephenate, PHE = phenylalanine, TYR = tyrosine, and PTN = "protein.". Enzymes: HK = hexokinase, PFK = phosphofructokinase, GAPD = glyceraldehyde-3-phosphate dehydrogenase, PK = pyruvate kinase, and AK = adenylate kinase. Pathway segments in amino acid biosynthesis are identified by numbers above reaction arrows. The flux through segment i is r^i. μ = specific growth rate for cell mass. A detailed discussion on the choice of individual components of this model can be found in Schlosser and Bailey [81] and Galazzo and Bailey [17, 18].

In the model which follows, all concentrations are in (mmol/L_{cells}) and rates are in (mmol/L_{cells}/min). Rate constants and other kinetic parameters in the equations have appropriate units. K_{eq} for the reaction ATP + AMP $\rightleftarrows$ 2ADP had a value of one. The total adenine nucleotide concentration was taken to be [ATP] + [ADP] + [AMP] = 3mM. These two relationships are used to solve for [ADP] and [AMP], given a value of [ATP]. The rate equations for the other reactions in the glycolytic pathway are as follows:

$$r_{in} = 200 - 132.5[\mathrm{G6P}] \tag{D.1}$$

$$r_{\mathrm{HK}} = 68.5\left(\frac{0.00062}{[\mathrm{G}][\mathrm{ATP}]} + \frac{0.11}{[\mathrm{G}]} + \frac{0.1}{[\mathrm{ATP}]} + 1\right)^{-1} \tag{D.2}$$

$$r_{\mathrm{POL}} = 15.74\left(\frac{[\mathrm{G6P}]^{8.51}}{93.0 + [\mathrm{G6P}]^{8.51}}\right)\left(\frac{2.558}{[\mathrm{G6P}]^2} + \frac{2.326}{[\mathrm{G6P}]} + 1\right)^{-1} \tag{D.3}$$

$$r_{\mathrm{PFK}} = \frac{3019[\mathrm{G6P}][\mathrm{ATP}]R}{R^2 + 6253L^2T^2} \tag{D.4}$$

$$R = 1 + 0.5714[\mathrm{G6P}] + 16.67[\mathrm{ATP}] + 95.24[\mathrm{G6P}][\mathrm{ATP}] \tag{D.5}$$

$$T = 1 + 0.0002857[\mathrm{G6P}] + 16.67[\mathrm{ATP}] + 0.004762[\mathrm{G6P}][\mathrm{ATP}] \tag{D.6}$$

$$L = \frac{1 + 0.76[\mathrm{AMP}]}{1 + 40[\mathrm{AMP}]} \tag{D.7}$$

$$r_{\mathrm{GAPD}} = 49.8\left(1 + \frac{0.25}{[\mathrm{FDP}]}\right. \tag{D.8}$$

$$\left. + \left(0.09375 + \frac{6.273}{[\mathrm{FDP}]}\right)(1 + 0.9091[\mathrm{ATP}] + 0.6667[\mathrm{ADP}] + 0.4[\mathrm{ATP}])\right)^{-1}$$

$$r_{\mathrm{PK}} = \frac{9763[\mathrm{PEP}][\mathrm{ADP}](R_{\mathrm{PK}} + 0.3964 L_{\mathrm{PK}}^2 T_{\mathrm{PK}})}{R_{\mathrm{PK}}^2 + 311.2\ L_{\mathrm{PK}}^2\ T_{\mathrm{PK}}^2} \tag{D.9}$$

$$R_{\mathrm{PK}} = 1 + 157.0[\mathrm{PEP}] + 0.2[\mathrm{ADP}] + 3.140[\mathrm{PEP}][\mathrm{ADP}] \tag{D.10}$$

$$T_{\mathrm{PK}} = 1 + 0.02[\mathrm{PEP}] + 0.2[\mathrm{ADP}] + 0.004[\mathrm{PEP}][\mathrm{ADP}] \tag{D.11}$$

$$L_{\mathrm{PK}} = \frac{1 + 0.05[\mathrm{FDP}]}{1 + 5[\mathrm{FDP}]} \tag{D.12}$$

$$r_{\mathrm{ATPase}} = 12.1[\mathrm{ATP}]. \tag{D.13}$$

The rate equations for the steps in the aromatic amino acid pathway are as follows:

$$r_1 = 4568 \left(\frac{\frac{0.79}{1+[\mathrm{PHE}]/53} + \frac{0.2}{1+[\mathrm{TYR}]/40} + \frac{0.01}{1+[\mathrm{TRP}]/16}}{\left(\frac{0.0002}{[\mathrm{PEP}][\mathrm{G6P}]} + \frac{0.006}{[\mathrm{PEP}]}\right)(1 + 50[\mathrm{DAHP}]) + \frac{0.1}{[\mathrm{G6P}]} + 1} \right) \tag{D.14}$$

$$r_2 = \frac{116.0[\mathrm{DAHP}][\mathrm{PEP}][\mathrm{ATP}]}{(2 + [\mathrm{DAHP}])(0.008665 + [\mathrm{PEP}])(0.9281 + [\mathrm{ATP}])} \tag{D.15}$$

$$r_3 = \frac{475.4[\mathrm{CHR}]}{(2 + [\mathrm{CHR}])(1 + [\mathrm{PHE}]/50)(1 + [\mathrm{TYR}]/40)} \tag{D.16}$$

$$r_4 = \frac{63.64[\mathrm{PHP}]}{1 + [\mathrm{PHP}](1 + [\mathrm{PHE}]/50)} \tag{D.17}$$

$$r_5 = \frac{10.48[\mathrm{PHP}]}{1 + [\mathrm{PHP}]} \tag{D.18}$$

$$r_6 = \frac{75.6[\mathrm{G6P}][\mathrm{CHR}][\mathrm{ATP}]}{(1.269 + [\mathrm{G6P}])(2 + [\mathrm{CHR}])(0.9281 + [\mathrm{ATP}])(1 + [\mathrm{TRP}]/16)}. \tag{D.19}$$

The rates at which the amino acids are utilized in the growing cell mass are taken to be proportional to the specific growth rate, μ, with 176mM, 131mM and 54mM utilized per unit cell volume for PHE, TYR and TRP, respectively. Finally, a rate expression for the specific growth rate was selected to be

$$\mu = \frac{0.034 Y\ [\mathrm{PHE}][\mathrm{TYR}][\mathrm{TRP}][\mathrm{PEP}]}{0.25 + Y\ (18 + [\mathrm{PHE}])(13 + [\mathrm{TYR}])(5 + [\mathrm{TRP}])(0.005923 + [\mathrm{PEP}])} \tag{D.20}$$

where

$$Y = \frac{[\mathrm{ATP}] + \frac{1}{2}[\mathrm{ADP}]}{[\mathrm{ATP}] + [\mathrm{ADP}] + [\mathrm{AMP}]}. \tag{D.21}$$

Amino acid uptake is assumed to be limited by its supply in the medium, and

hence is an adjustable parameter set by the "experimenter." f^0_{PHE}, f^0_{TYR} and f^0_{TRP} are the rates at which PHE, TYR and TRP are fed, respectively. The mass balance equations for each of the cellular constituents are then given by:

$$\mathrm{G}: \quad 0 = r_{in} - r_{\mathrm{HK}} - \mu[\mathrm{G}] \tag{D.22}$$

$$\mathrm{G6P}: \quad 0 = r_{\mathrm{HK}} - r_{\mathrm{POL}} - r_{\mathrm{PFK}} - r_1 - r_6 - \mu[\mathrm{G6P}] \tag{D.23}$$

$$\mathrm{FDP}: \quad 0 = r_{\mathrm{PFK}} - r_{\mathrm{GAPD}} - \mu[\mathrm{FDP}] \tag{D.24}$$

$$\mathrm{PEP}: \quad 0 = 2r_{\mathrm{GAPD}} - r_{\mathrm{PK}} - r_1 - r_2 - \mu[\mathrm{PEP}] \tag{D.25}$$

$$\mathrm{ATP}: \quad 0 = -r_{\mathrm{HK}} - r_{\mathrm{POL}} - r_{\mathrm{PFK}} + 2r_{\mathrm{GAPD}} + r_{\mathrm{PK}} - r_{\mathrm{ATPas}} - r_2 - 4r_6 - \mu[\mathrm{ATP}] \tag{D.26}$$

$$\mathrm{DAHP}: \quad 0 = r_1 - r_2 - \mu[\mathrm{DAHP}] \tag{D.27}$$

$$\mathrm{CHR}: \quad 0 = r_2 - r_3 - r_6 - \mu[\mathrm{CHR}] \tag{D.28}$$

$$\mathrm{PHP}: \quad 0 = r_3 - r_4 - r_5 - \mu[PHP] \tag{D.29}$$

$$\mathrm{PHE}: \quad 0 = r_4 + f^0_{\mathrm{PHE}} - \mu(176 + [\mathrm{PHE}]) \tag{D.30}$$

$$\mathrm{TYR}: \quad 0 = r_5 + f^0_{\mathrm{TYR}} - \mu(131 + [\mathrm{TYR}]) \tag{D.31}$$

$$\mathrm{TRP}: \quad 0 = r_6 + f^0_{\mathrm{TRP}} - \mu(54 + [\mathrm{TRP}]). \tag{D.32}$$

D.2 Black Frame Analysis of the Cell Metabolism Model

The following is a description of the calculations involved in applying a Black Frame analysis [81] to the model described in the prior section where the "cellular metabolic environment" represents the Black Frame or subset of metabolism. This "Black Frame" analysis is needed to estimate an elasticity and to complete the analysis. In particular, the calculations presented here allow one to obtain concentration control

coefficients (*cccs*) for this system, given that one knows the rate equations for all steps in the "detailed metabolic pathway" of Figure 11.1 and has obtained a value for the elasticity of the Black Frame without knowing any of the related details. First, some definitions:

$$cx = \begin{bmatrix} [\text{DAHP}] \\ [\text{CHR}] \\ [\text{PHP}] \\ [\text{TRP}] \\ [\text{TYR}] \\ [\text{PHE}] \end{bmatrix}, \quad cy = \begin{bmatrix} [\text{G6P}] \\ [\text{PEP}] \\ [\text{ATP}] \end{bmatrix}, \quad f = \begin{bmatrix} \mu[\text{DAHP}] \\ \mu[\text{CHR}] \\ \mu[\text{PHP}] \\ \mu([\text{TRP}] + 176) \\ \mu([\text{TYR}] + 131) \\ \mu([\text{PHE}] + 54) \end{bmatrix}, \quad v = \begin{bmatrix} r_1 \\ r_2 \\ r_3 \\ r_4 \\ r_5 \\ r_6 \end{bmatrix},$$

$$K = \begin{bmatrix} 1 & -1 & 0 & 0 & 0 & 0 \\ 0 & 1 & -1 & 0 & 0 & -1 \\ 0 & 0 & 1 & -1 & -1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}, \quad y = \begin{bmatrix} [\text{G6P}] \\ [\text{PEP}] \\ [\text{ATP}] \\ \mu \end{bmatrix}, \quad p = [\mu], \text{ and } x = \begin{bmatrix} [\text{TRP}] \\ [\text{TYR}] \\ [\text{PHE}] \\ r_1 \\ r_2 \\ r_6 \end{bmatrix}.$$

After performing a series of experiments to determine the steady-state values of x and y under various perturbations, as well as for the undisturbed system, one can form the matrices X and Y as described in the main text and calculate an elasticity, R, such that $Y \approx XR + E$. R can then be broken into matrices H_1, H_2, G_1, and G_2, filling in zeros where appropriate, such that

$$\Delta cy \approx H_1 \Delta cx + H_2 \Delta v \quad \text{and} \tag{D.33}$$

$$\Delta p \approx G_1 \Delta cx + G_2 \Delta v. \tag{D.34}$$

Here, Δ means the difference between the steady-state value under disturbance and

the undisturbed steady-state value. One then calculates the values of $\nabla_{cx}v$, $\nabla_{cy}v$, $\nabla_{cx}f$, and $\nabla_{p}f$ from the model equations via numerical differentiation. Then

$$J_1 = (K + \nabla_p f \cdot G_2) \cdot \left(\nabla_{cx} v + \nabla_{cy} v \cdot (I - H_2 \cdot \nabla_{cy} v)^{-1} \cdot (H_1 + H_2 \cdot \nabla_{cx} v)\right) + \nabla_{cx} f + \nabla_p f \cdot G_1, \quad \text{(D.35)}$$

$$J_2 = (K + \nabla_p f \cdot G_2) \cdot \left(I + \nabla_{cy} v \cdot (I - H_2 \cdot \nabla_{cy} v)^{-1} \cdot H_2\right), \quad \text{(D.36)}$$

and the concentration control coefficient ccc of the jth step for the ith ouptut is

$$(ccc)_{ij} \equiv \left(\frac{d\ cx_i}{d\ r_j^{max}}\right) \cdot \left(\frac{r_j^{max}}{(cx_i)_0}\right) = \left(\frac{(r_j)_0}{(cx_i)_0}\right) \cdot \left(-J_1^{-1} \cdot J_2\right)_{ij} \quad \text{(D.37)}$$

where the 0 subscripts indicate the undisturbed steady-state values.

D.3 Perturbations Used with the Cellular Metabolism Model

Table D.1 describes the perturbations studied by Galazzo and Bailey [17, 18]. These peturbations did not have a high enough SNR for any of the studied regressors to produce useful cccs. Feed rates for PHE, TYR and TRP, corresponding to f^0_{PHE}, f^0_{TYR} and f^0_{TRP} respectively, are expressed relative to the unperturbed fluxes to those species, $(r_4)_0$, $(r_5)_0$ and $(r_6)_0$. For example, in Experiment 1, one would set $f^0_{\text{PHE}} = 0.05(r_4)_0$, and $f^0_{\text{TYR}} = f^0_{\text{TRP}} = 0$. Inhibitions of r_1 are obtained by multiplying r_1^{max} by the appropriate fraction; e.g., in Experiment 7, r_1^{max} would become $0.9 \times 4568 = 4111.2$.

The experiemental perturbations used in this simulation study are described in Table 11.1. Experimental deviations are the differences between the steady-state values of the model under perturbation and the unperturbed steady-state values.

Experiment No.	Feed Rates			Inhibition of r_1
	PHE	TYR	TRP	
1	5%	-	-	-
2	10%	-	-	-
3	-	5%	-	-
4	-	10%	-	-
5	-	-	5%	-
6	-	-	10%	-
7	-	-	-	10%
8	-	-	-	20%

Table D.1. Experimental perturbations used in Schlosser and Bailey [81]

Changes in [G6P] were all below "detectable" levels.

Quantity	Nominal Value	Experiment No.						
		1	2	3	4	5	6	7
[TRP]	80.0	32.0	1.0	6.0	2.0	33.0	7.0	11.0
[TYR]	316.0	-2.0	15.0	-17.0	-22.0	-23.0	103.0	-28.0
[PHE]	261.0	-1.0	27.0	72.0	-9.0	-10.0	-30.0	65.0
r_1	9.25	0.13	0.00	-0.15	-0.12	-0.04	-0.15	-0.35
r_2	9.22	0.14	-0.01	-0.15	-0.11	-0.03	-0.14	-0.34
r_6	1.18	-0.14	0.00	0.10	0.05	-0.10	0.10	0.19
[PEP]	0.014	-0.002	0.001	-0.001	-0.002	-0.003	-0.001	-0.002
[ATP]	0.24	0.05	-0.02	0.03	0.06	0.11	0.03	0.09
μ	0.534	0.018	0.0	0.018	0.018	0.03	0.018	0.042

Table D.2. Nominal and deviation values "measured" from the nonlinear model for the perturbations in Table 11.1

Bibliography

[1] J. Berkson. Are there two regressions? *American Statistical Association Journal*, 45:164–180, 1950.

[2] J. Bibby and H. Toutenburg. *Prediction and Improved Estimation in Linear Estimators*. Wiley, 1977.

[3] W. E. Biles and J. J. Swain. *Optimization and Industrial Experimentation*. Wiley, 1980.

[4] G. E. P. Box, W. G. Hunter, and J. S. Hunter. *Statistics for Experimenters*. Wiley, 1978.

[5] C. Brosilow and B. Joseph. Inferential control of processes. *AIChE Journal*, 24:485–509, 1978.

[6] G. C. Brown, R. P. Hafner, and M. D. Brand. A "top-down" approach to the determination of control coefficients in metabolic control theory. *Eur. J. Biochem.*, 188:321–325, 1990.

[7] H. M. Budman, T. R. Holcomb, and M. Morari. PLS-based robust inferential control for a packed-bed reactor. In *Proceedings of the 1991 Automatic Control Conference*, pages 256–261, 1991.

[8] H. M. Budman, C. Webb, T. R. Holcomb, and M. Morari. Robust inferential control for a packed-bed reactor. *Industrial and Chemical Engineering Research*, 37(7):1665–1679, 1992.

[9] E. I. Canela, M. Cascante, and R. Franco. In *Control of Metabolic Processes*, volume 188, pages 157–169. Plenum Press, New York, 1990.

[10] W. P. Carey, K. R. Beebe, E. Sanchez, and B. Kowalski. Chemometric analysis of multisensor arrays. *Sensors and Actuators*, 9:223–234, 1986.

[11] G. C. Chow. *Econometrics*. McGraw-Hill, 1983.

[12] A. P. Dempster, N. M. Laird, and D. B. Rubin. Maximum likelihood from incomplete data via the E-M algorithm. *Journal of the Royal Statistical Society*, B39:1–22, 1977.

[13] J. Doyle. Analysis of feedback systems with structured uncertainties. *IEEE Proceedings Part D*, 129:242–250, 1982.

[14] N. R. Draper and H. Smith. *Applied Regression Analysis.* Wiley, 1966.

[15] T. F. Edgar and D. M. Himmelblau. *Optimization of Chemical Processes.* McGraw-Hill, 1988.

[16] D. Fell and H. M. Sauro. Metabolic control theory and its analysis: additional relationships between elasticities and control coefficients. *Eur. J. Biochem*, 148:555–561, 1985.

[17] J. Galazzo and J. Bailey. Fermentation pathway kinetics and metabolic flux control in suspended and immobilized saccharomyces cerevisiae. *Enzyme Microb. Technol*, 12:162–172, 1990.

[18] J. Galazzo and J. Bailey. Errata. *Enzyme Microb. Technol*, 13:363, 1991.

[19] K. F. Gauss. *Theoria motus corporum celestium.* English Translation: *Theory of the Motion of the Heavenly Bodies.* Dover (1963), 1809.

[20] P. Geladi and B. Kowalski. Partial least squares regression: A tutorial. *Analytica Chimica Acta*, 185:1–17, 1986.

[21] G. H. Golub and C. F. V. Loan. *Matrix Computations.* John Hopkins, second edition, 1989.

[22] C. Goodall. *Understanding Robust and Exploratory Data Analysis*, chapter *M*-Estimators of Location: An Outline of the Theory. Wiley, 1983.

[23] S. J. Gould. *The Mismeasure of Man.* Norton, 1981.

[24] M. H. Gruber. *Regression Estimators.* Academic Press, 1990.

[25] T. Haitovsky and Y. Wax. Generalized ridge regression, least squares with stochastic prior information, and Bayesian estimators. *Applied Mathematics and Computation*, 7:125–154, 1980.

[26] F. R. Hampel. *Contributions to the Theory of Robust Estimation.* Ph.D . thesis, University of California, Berkeley, 1968.

[27] T. Harris, J. MacGregor, and J. Wright. Optimal sensor location with an application to a packed bed tubular reactor. *AIChE Journal*, 26, 1980.

[28] R. Heinrich and T. Rapoport. A linear steady-state treatment of enzymatic chains. *Eur. J. Biochem.*, 42:89–95, 1974.

[29] I. S. Helland. On the structure of partial least squares. *Communications in Statistics - Simulation*, 17(2):1581–607, 1988.

[30] I. S. Helland. Partial least squares and statistical models. *Scandinavian Journal of Statistics*, 17:97–114, 1990.

[31] I. S. Helland. Maximum likelihood regression on relevant components. *Journal of the Royal Statistial Society, Series B: Methodological*, 54(2):637–647, 1992.

[32] R. C. Hill, T. B. Fomby, and S. Johnson. Component selection norms for principal components regression. *Communications in Statistics A: Theory and Methods*, A6(4):309–334, 1977.

[33] D. C. Hoaglin, F. Mosteller, and J. W. Tukey. *Understanding Robust and Exploratory Data Analysis.* Wiley, 1983.

[34] A. E. Hoerl and R. W. Kennard. Ridge regression: Biased estimation for nonorthogonal problems. *Technometrics*, 12:55–67, 1970.

[35] T. R. Holcomb and M. Morari. PLS leads to different algorithms for factor analysis and regression. in preparation, 1992.

[36] T. R. Holcomb and M. Morari. PLS/neural networks. *Computers and Chemical Engineering*, 16(4):393–411, 1992.

[37] B. Holt and M. Morari. Design of resilient processing plants V: the effect of deadtime on dynamic resilience. *Chemical Engineering Science*, 40:1229–1237, 1985.

[38] R. A. Horn and C. R. Johnson. *Matrix Analysis.* Cambridge, 1985.

[39] A. Höskuldsson. PLS regression methods. *Journal of Chemometrics*, 2:211–228, 1988.

[40] P. J. Huber. Robust estimation of a location parameter. *Annals of Statistics*, 35:73–101, 1964.

[41] P. J. Huber. The behavior of maximum likelihood estimates under nonstandard conditions. In *Fifth Berkeley Symposium on Mathematical Statistics and Probability*, volume 1, pages 73–101, 1967.

[42] P. J. Huber. *Robust Statistical Procedures.* SIAM, 1977.

[43] L. Jancar and W. Wegscheider. Effect of scaling regimes on the prediction of analytical results from multivariate calibration. *Analytica Chimica Acta*, pages 459–472, 1991.

[44] J. Johnston. *Econometric Methods, 2nd Edition.* McGraw-Hill, 1972.

[45] I. T. Joliffe. A note on the use of principal components in regression. *Applied Statistics*, 31:300–303, 1982.

[46] I. T. Joliffe. *response to*: Continuum regression: Cross-validated sequentially constructed prediction embracing ordinary least squares, partial least squares, and principal components regression. *Journal of the Royal Statistical Society*, B52:266, 1990.

[47] S. Jørgensen, L. Goldschmidt, and K. Clement. A sensor-location procedure for chemical processes. *Computers and Chemical Engineering*, 8:195–204, 1984.

[48] H. Kacser and J. Burns. The control of flux. *Symp. Soc. Exp. Biol.*, 27:65–104, 1973.

[49] D. Kell and H. Westerhoff. Metabolic control theory: its role in microbiology and biotechnology. *FEMS Microbiol. Rev.*, 39:305–320, 1986.

[50] D. Kell and H. Westerhoff. Matrix method for determining steps most rate-limiting to metabolic fluxes in biotechnological processes. *Biotechnol. Bioeng.*, 30:101–107, 1987.

[51] J. Kresta, J. MacGregor, and T. E. Marlin. Multivariate statistical monitoring of process operating performance. *Canadian Journal of Chemical Engineering*, 69(1):35–47, 1991.

[52] S. Kumar and J. Seinfeld. Optimal locations of measurements in tubular reactors. *Chemical Engineering Science*, 33:1507–1516, 1978.

[53] D. Laughlin, K. Jordan, and M. Morari. Internal model control and process uncertainty - mapping uncertainty regions for SISO controller design. *International Journal of Control*, 44:1675–1698, 1986.

[54] J. Lee and M. Morari. Robust measurement selection. *Automatica*, 27(3):519–527, 1991.

[55] D. V. Lindley. Regression lines and the linear functional relationship. *Journal of the Royal Statistical Society*, B(Supplement) 52:218–244, 1947.

[56] L. Ljung. *System Identification: Theory for the User.* Prentice Hall, 1987.

[57] A. Lorber and B. Kowalski. Alternatives to cross-validatory estimation of the number of factors in multivariate calibration. *Applied Spectroscopy*, 44(9):1464–1470, 1990.

[58] A. Lorber, L. Wangen, and B. Kowalski. A theoretical foundation for the PLS algorithm. *Journal of Chemometrics*, 5:19–31, 1987.

[59] D. J. MacKay. *Bayesian methods for adaptive models.* Ph.D. thesis, California Institute of Technology, 1992.

[60] J. A. Mandler. *Robust Control System Design for a Fixed-Bed Catalytic Reactor.* Ph.D. thesis, California Institute of Technology, 1987.

[61] D. W. Marquardt and R. D. Snee. Ridge regression in practice. *The American Statistician*, 29:3–20, 1975.

[62] H. Martens and T. Næs. *Multivariate Calibration.* Wiley, 1989.

[63] T. Mejdell. *Estimators for Product Composition in Distillation Columns.* Ph.D. thesis, University of Trondheim, The Norway Institute of Technology, November 1990.

[64] C. Moler, J. Little, S. Bangert, and S. Kleinman. *MATLAB User's Guide.* The MathWorks, 1990.

[65] P. Moran. Estimating structural and functional relationships. *Journal of Multivariate Analysis*, 1:232–255, 1971.

[66] M. Morari and E. Zafiriou. *Robust Process Control.* Prentice Hall, 1989.

[67] T. Næs. *response to*: Continuum regression: Cross-validated sequentially constructed prediction embracing ordinary least squares, partial least squares, and principal components regression. *Journal of the Royal Statistical Society*, B.52:263, 1990.

[68] A. Negiz and A. Cinar. On the detection of multiple sensor abnormalities in multivariate processes. In *Proceedings of the 1992 Automatic Control Conference*, pages 2364–2368, 1992.

[69] I. Olkin, editor. *Contributions to Probability and Statistics*, A survey of sampling from contaminated distributions. Stanford University Press, 1960.

[70] L. Peele and T. P. Ryan. Minimax linear estimators with applications to ridge regression. *Technometrics*, 24(2):157–159, 1986.

[71] M. Piovoso, K. Kosanovich, and R. K. Pearson. Monitoring process performance in real-time. In *Proceedings of the 1992 Automatic Control Conference*, pages 2359–2364, 1992.

[72] M. J. D. Powell. Radial basis functions for multivariable interpolation: A review. In *Algorithms for Approximation*, pages 143–167. Clarendon Press, Oxford, 1987.

[73] S. Qin and T. McAvoy. Nonlinear PLS modeling using neural networks. *Computers and Chemical Engineering*, 16(4):379–391, 1992.

[74] C. Reder. Metabolic control theory: a structural approach. *J. Theor. Biol.*, 135:175–201, 1988.

[75] N. L. Ricker. The use of biased least-squares estimators for parameters in discrete-time pulse response models. *Industrial and Engineering Chemical Research*, 27:343–350, 1988.

[76] S. M. Ross. *Introduction to Probability and Statistics for Engineers and Scientists.* Wiley, 1987.

[77] E. Sanchez and B. Kowalski. Generalized rank annihilation factor analysis. *Analytical Chemistry*, 58(2):496–499, 1986.

[78] M. A. Savageau and A. Sorribas. Strategies for representing metabolic pathways within biochemical systems theory: reversible pathways. *Math. Biosci.*, 94:239–269, 1989.

[79] M. A. Savageau and E. O. Voit. Power-law approach to modelling biological systems: I. theory. *J. Ferment. Technol.*, 60:221–228, 1982.

[80] P. Schlosser. Personal communication, 1992.

[81] P. Schlosser and J. Bailey. An integrated modelling-experimental strategy for the analysis of metabolic pathways. *Math. Biosci*, 100:87–114, 1990.

[82] P. M. Schlosser, T. R. Holcomb, and J. E. Bailey. Determining sensitivity coefficients directly from experimental data. *Biotechnology and Bioengineering*, accepted 1992.

[83] H. Schneeweiss. Consistent estimation of a regression with errors in the variables. *Metrika*, 23:101–115, 1976.

[84] S. D. Silvey. Multicollinearity and imprecise estimation. *Journal of the Royal Statistical Society*, B.31:539–552, 1969.

[85] G. Smith and F. Campbell. A critique of some ridge regression methods. *Journal of American Statistical Association*, 75(369):74–100, March 1980.

[86] P. Sprent. Some history of functional and structural relationships. In *Statistical Analysis of Measurement Error Models and Applications*, pages 3–15. American Mathematical Society, American Mathematical Society, 1990.

[87] M. Stone. Cross-validatory choice and assessment of statistical predictions (with discussion). *Journal of the Royal Statistical Society*, B.36:111–147, 1974.

[88] M. Stone and R. J. Brooks. Continuum regression: Cross-validated sequentially constructed prediction embracing ordinary least squares, partial least squares, and principal components regression. *Journal of the Royal Statistical Society*, B.52:237–269, 1990.

[89] G. Strang. *Linear Algebra and Its Applications.* Academic Press, second edition, 1980.

[90] M. T. Subrahmanya. A note on non-negative estimators of postive parameters. *Metrika*, 16:107–114, 1970.

[91] H. Theil and A. S. Goldberger. On pure and mixed estimation in economics. *International Economic Review*, 2:65–78, 1961.

[92] C. M. Theobald. Generalizations of mean squared error applied to ridge regression. *Journal of the Royal Statistical Society*, B.36:103–106, 1974.

[93] J. Tobin. Estimation of relationships for limited dependent variables. *Econometrica*, 26:24–36, 1958.

[94] B. Wahlberg and L. Ljung. Hard frequency-domain model error bounds from least-squares like identification techniques. *IEEE Transactions on Automatic Control*, 37(7), 1992.

[95] C. Webb. *Robust Control Strategies for a Fixed Bed Chemical Reactor.* Ph.D. thesis, California Institute of Technology, 1990.

[96] C. Webb, H. Budman, and M. Morari. Identifying frequency domain uncertainty bounds for robust controller design – theory with applications to a fixed bed reactor. In *Proceedings of the American Control Conference*, pages 1528–1533, 1989.

[97] C. Webb, H. Budman, and M. Morari. Robust inferential control for a packed bed reactor. In *Proceedings of the American Control Conference*, pages 575–580, 1990.

[98] S. Wold. Cross-validatory estimation of the number of components in factor and principal components models. *Technometrics*, 20(4):397–405, 1978.

[99] S. Wold, P. Geladi, K. Esbensen, and J. Öhman. Multi-way principal components- and PLS-analysis. *Chemometrics and Intelligent Laboratory Systems*, 1:41–56, 1987.

[100] S. Wold, N. Kettaneh-Wold, and B. Skagerberg. Nonlinear PLS modeling. *Chemometrics and Intelligent Laboratory Systems*, 7(1-2):53–65, December 1989.

[101] S. Wold, A. Ruhe, H. Wold, and W. Dunn. The collinearity problem in linear regression: The partial least squares approach to generalized inverses. *SIAM J. Sci. Stat. Comput.*, 5(3):753–743, September 1984.

[102] Y. D. Wang, D. J. Veltkamp, and B. R. Kowalski. Multivariate instrument standardization. *Analytical Chemistry*, 63(23):2750–2756, 1991.