University of Southampton

Faculty of Engineering, Science and Mathematics

School of Mathematics

Simultaneous Confidence Bands In Linear Regression Analysis

by

Pascal Soon Shien Ah-kine

Thesis for the degree of Doctor of Philosophy

December 11, 2009

Abstract

A simultaneous confidence band provides useful information on the plausible range of an unknown regression model. For a simple linear regression model, the most frequently quoted bands in the statistical literature include the two-segment band, the three-segment band and the hyperbolic band, and for a multiple linear regression model, the most common bands in the statistical literature include the hyperbolic band and the constant width band. The optimality criteria for confidence bands include the Average Width criterion considered by Gafarian (1964) and Naiman (1984) among others, and the Minimum Area Confidence Set (MACS) criterion of Liu and Hayter (2007). A concise review of the construction of two-sided simultaneous confidence bands in simple and multiple linear regressions and their comparison under the two mentioned optimality criteria is provided in the thesis. Two families of confidence bands, the *inner-hyperbolic* bands and the *outer-hyperbolic* bands, which include the hyperbolic and three-segment bands as special cases, are introduced for a simple linear regression. Under the MACS criterion, the best confidence band within each family is found by numerical search and compared with the hyperbolic band, the best three-segment band and with each other. The *inner-hyperbolic* family of confidence bands, which include the hyperbolic and constant-width bands as special cases, is also constructed for a multiple linear regression model over an ellipsoidal covariate region and the best band within the family is found by numerical search. For a multiple linear regression model over a rectangular covariate region (*i.e.* the predictor variables are constrained in intervals), no method of constructing exact simultaneous confidence bands has been published so far. A method to construct exact two-sided hyperbolic and constant width bands over a rectangular covariate region and compare between them is provided in this thesis when there are up to three predictor variables. A simulation method similar to the ones used by Liu *et al.* (2005a) and Liu *et al.* (2005b) is also provided for the calculation of the average width and the minimum volume of confidence set when there are more than three predictor variables. The methods used in this thesis are illustrated with numerical examples and the Matlab programs used are available upon request.

Contents

List of Figures

List of Tables

Declaration Of Authorship

I, Pascal Soon Shien Ah-kine, declare that the thesis entitled Simultaneous Confidence Bands In Linear Regression Analysis and the work presented in the thesis are both my own, and have been generated by me as the result of my own original research. I confirm that:

- this work was done wholly or mainly while in candidature for a research degree at this University;
- where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated;
- where I have consulted the published work of others, this is always clearly attributed;
- where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work;
- I have acknowledged all main sources of help;
- where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself;
- none of this work has been published before submission. Parts of this work are intended to be submitted for publication as:
  - Liu, W. and Ah-Kine, P. Optimal Simultaneous Confidence Bands in Simple Linear Regression
  - Ah-Kine, P. and Liu, W. Optimal Simultaneous Confidence Bands in Multiple Linear Regression with Predictor Variables Constrained in an Ellipsoidal Region.
  - Liu, W. and Ah-Kine, P. Exact simultaneous confidence bands for a linear regression with predictor variables constrained in intervals.

Signed: .....

Date : .....

Acknowledgements

Foremost, I would like to express my utmost gratitude to the School of Mathematics of the University of Southampton and the Overseas Research Students Awards Scheme (ORSAS) for funding my research. Not only was I given the honour to research in the statistical field but I was also given the opportunity to travel to various places to present my work.

This thesis would not have been possible without the caring supervision of Prof. Wei Liu. I am very grateful for his guidance, patience, motivation and immense knowledge he provided during the time of my research. Never short of brilliant ideas and insightful comments, Prof. Wei Liu has also given me the chance to meet and collaborate with other experts in the field of confidence bands. As a result, I have co-authored for the first time in the papers:

- Hayter, A.J., Liu, W. and Ah-Kine, P. (2009) A ray method of confidence band construction for multiple linear regression models. *Journal of Statistical Planning and Inference, 139, (2), 329-334.*

- Liu, W., Hayter, A.J., Piegorsch, W.W. and Ah-Kine, P. (2009). Comparison of hyperbolic and constant width simultaneous confidence bands in multiple linear regression under MVCS criterion. *Journal of Multivariate Analysis, 100, 1432-1439.*

Subsequently, I have been actively collaborating with Prof. Wei Liu in writing several other papers, three of which are related to this thesis. Suffice it to say that I could not have imagined having a better supervisor and mentor for my PhD study.

Besides my supervisor, I would also like to thank the other member of my thesis committee, Dr. Dave Woods. His teaching, support and presence over the years have been valuable, mainly by being the statistical figure that I look up to. I would also like to thank my undergraduate personal tutor, Dr. Sujit Sahu, who guided me to the statistical field and encouraged me to take admission in the PhD study. My sincere thanks also go to all the staff, colleagues and friends of the School of Mathematics and the Southampton Statistical Sciences Research Institute (S3RI), in particular Antony Overstall who has been sharing an office with me for the whole duration of my research.

I dedicate my thesis to my supportive family, especially my parents Vincent and Josiane Ah-kine, for the spiritual, moral and financial support throughout my life. Finally, I would like to give praise to my other half, Amy Lim, who has provided me with encouragement, inspiration and affection throughout my PhD study.

Chapter 1

Introduction to linear regression analysis and simultaneous confidence bands

The term "regression" originates from the 14^{th} century, where it had a biological meaning as "the act of going back". It was first adapted to a more general statistical context by the well-known statisticians Udny Yule and Karl Pearson. However, its first statistical form was published by Legendre (1805) and by Gauss (1809) in the field of astronomy, where they applied the method of least squares to the problem of determining orbits of bodies about the Sun. Since then, regression analysis has been widely applied to the study of biology, behavioral and social sciences and more recently in finance, industry and many other practical aspects of real life.

The first widely studied form of regression analysis has been linear regression, due to the simplicity of the model and the statistical properties of the estimators. Linear regression is usually used for the purpose of hypothesis testing or for the purpose of prediction and forecasting. Many statistical methods and techniques have emerged from its study and one of them is the simultaneous confidence band. This chapter provides a general review of linear regression and presents some preliminary results necessary for the construction and comparison of simultaneous confidence bands throughout the thesis.

1.1 Linear regression models

Linear regression analysis is a statistical technique used to model data consisting of a dependent random variable and one or more independent variables, so as to evaluate the relationship between the dependent variable and the independent variables. Specifically, the dependent variable y is expressed as a function of the independent variables $x_1, \ldots, x_k$, the corresponding parameters $b_0, b_1, \ldots, b_k$ and an error term e as in

$$y = b_0 + b_1 x_1 + b_2 x_2 + \ldots + b_k x_k + e. \tag{1.1}$$

The error term e is a random variable that represents the unexplained variation in the dependent variable y. If a sample of n observations are available with the i^{th} observation given by $(y_i, x_{i1}, x_{i2}, \ldots, x_{ik})$ for $i = 1, \ldots, n$, the i^{th} observation is assumed to satisfy the relationship

$$y_i = b_0 + b_1 x_{i1} + b_2 x_{i2} + \ldots + b_k x_{ik} + e_i$$

where $b_0, b_1, \ldots, b_k$ are the same for all observations. The linear regression model can also be represented in the matrix form

$$\boldsymbol{Y} = X\boldsymbol{b} + \boldsymbol{e} \tag{1.2}$$

$$\text{where } Y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} \qquad X = \begin{pmatrix} 1 & x_{11} & x_{12} & \cdots & x_{1k} \\ 1 & x_{21} & x_{22} & \cdots & x_{2k} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n1} & x_{n2} & \cdots & x_{nk} \end{pmatrix} \qquad \boldsymbol{b} = \begin{pmatrix} b_0 \\ b_1 \\ \vdots \\ b_k \end{pmatrix} \qquad \boldsymbol{e} = \begin{pmatrix} e_1 \\ e_2 \\ \vdots \\ e_n \end{pmatrix}$$

The matrix X is called the design matrix as its components can be suitably chosen via design. Moreover, the linear regression model is subject to the following assumptions:

- The errors follow a normal distribution with the mean zero and constant variance $\sigma^2 > 0$ and they are independent.
- The independent variables $x_1, \ldots, x_k$ are error-free and the design matrix X has full column rank $k + 1$.

1.2 Parameter estimation

1.2.1 Least squares estimates

From the sample of n observations, the linear regression model can be used to evaluate the relationship between the dependent variable y and the independent variables $x_1, \ldots, x_k$ by estimating the parameters of the model. These include the $k + 1$ coefficients $b_0, b_1, \ldots, b_k$ and the variance σ^2 of the error term e. In order that all the $k + 2$ parameters can be estimated from the sample data, there should be at least $n \geq k + 2$ observations. A common method of estimating the parameters is the method of least squares.

Let $\hat{\boldsymbol{b}} = (\hat{b}_0, \hat{b}_1, \ldots, \hat{b}_k)^T$ estimate $\boldsymbol{b} = (b_0, b_1, \ldots, b_k)^T$ and let $\hat{\sigma}^2$ estimate σ^2. Then, $\hat{\boldsymbol{b}}$ estimates $\boldsymbol{b}$ by minimizing the sum of squares $\sum_{i=1}^{n}(y_i - \boldsymbol{b}^T\boldsymbol{x}_i)^2$ where $\boldsymbol{x}_i = (1, x_{i1}, x_{i2}, \ldots, x_{ik})^T$. Note that

$$\begin{aligned} & \sum_{i=1}^{n}(y_i - \boldsymbol{b}^T\boldsymbol{x}_i)^2 \\ = \; & (\boldsymbol{Y} - X\boldsymbol{b})^T(\boldsymbol{Y} - X\boldsymbol{b}) = \|\boldsymbol{Y} - X\boldsymbol{b}\|^2 \\ = \; & \boldsymbol{Y}^T\boldsymbol{Y} - \boldsymbol{Y}^T X\boldsymbol{b} - \boldsymbol{b}^T X^T\boldsymbol{Y} + \boldsymbol{b}^T X^T X\boldsymbol{b} \\ = \; & \boldsymbol{Y}^T\boldsymbol{Y} - 2\boldsymbol{b}^T X^T\boldsymbol{Y} + \boldsymbol{b}^T X^T X\boldsymbol{b} \end{aligned}$$

since $\boldsymbol{Y}^T X\boldsymbol{b}$ is a scalar and its transpose $\boldsymbol{b}^T X^T\boldsymbol{Y}$ is the same scalar. Now

$$\begin{aligned} & \frac{\partial}{\partial \boldsymbol{b}}(\boldsymbol{Y}^T\boldsymbol{Y} - 2\boldsymbol{b}^T X^T\boldsymbol{Y} + \boldsymbol{b}^T X^T X\boldsymbol{b}) \\ = \; & -2X^T\boldsymbol{Y} + 2X^T X\boldsymbol{b} \\ = \; & 0 \;\text{ at }\; \boldsymbol{b} = \hat{\boldsymbol{b}}. \end{aligned}$$

Hence, $X^T X\hat{\boldsymbol{b}} = X^T\boldsymbol{Y}$ and if X has full column rank, then $X^T X$ is non-singular and so

$$\hat{\boldsymbol{b}} = (X^T X)^{-1} X^T\boldsymbol{Y}. \tag{1.3}$$

Since $e_1, e_2, \ldots, e_n$ are identically and independently distributed as $N(0, \sigma^2)$, $\hat{\boldsymbol{b}}$ is also the maximum likelihood estimator of $\boldsymbol{b}$. The fitted values for $\boldsymbol{Y}$ are $\hat{\boldsymbol{Y}} = X\hat{\boldsymbol{b}}$, the residuals are given by

$$\begin{aligned} \hat{\boldsymbol{e}} &= (\hat{e}_1, \hat{e}_2, \ldots, \hat{e}_n)^T \\ &= \boldsymbol{Y} - \hat{\boldsymbol{Y}} \\ &= \boldsymbol{Y} - X\hat{\boldsymbol{b}}, \end{aligned}$$

and

$$\begin{aligned} \hat{\sigma}^2 &= (\text{residual sum of squares}) \;/\; (n-k-1) \\ &= \sum_{i=1}^{n} (\hat{e}_i)^2 \;/\; (n-k-1) \\ &= \hat{\boldsymbol{e}}^T\hat{\boldsymbol{e}} \;/\; (n-k-1) \\ &= (\boldsymbol{Y} - X\hat{\boldsymbol{b}})^T(\boldsymbol{Y} - X\hat{\boldsymbol{b}}) \;/\; (n-k-1) \\ &= \frac{\|\boldsymbol{Y} - X\hat{\boldsymbol{b}}\|^2}{(n-k-1)}. \end{aligned} \tag{1.4}$$

1.2.2 Distributions of $\hat{\boldsymbol{b}}$, $\hat{\boldsymbol{e}}$ and $\hat{\sigma}^2$

Note that $\hat{\boldsymbol{b}} = (X^T X)^{-1} X^T\boldsymbol{Y}$, that is $\hat{\boldsymbol{b}}$ is a linear transformation of $\boldsymbol{Y}$, and $\boldsymbol{Y} \sim N(X\boldsymbol{b}, \sigma^2)$. Therefore, $\hat{\boldsymbol{b}}$ follows a normal distribution with mean

$$\begin{aligned} E(\hat{\boldsymbol{b}}) &= E\left[(X^T X)^{-1} X^T\boldsymbol{Y}\right] \\ &= (X^T X)^{-1} X^T E(\boldsymbol{Y}) \\ &= (X^T X)^{-1} X^T X\boldsymbol{b} \\ &= \boldsymbol{b} \end{aligned}$$

and variance

$$\begin{aligned} Var(\hat{\boldsymbol{b}}) &= Cov\left[(X^T X)^{-1} X^T\boldsymbol{Y},\; (X^T X)^{-1} X^T\boldsymbol{Y}\right] \\ &= (X^T X)^{-1} X^T Cov(\boldsymbol{Y},\; \boldsymbol{Y}) X (X^T X)^{-1} \\ &= \sigma^2 (X^T X)^{-1}. \end{aligned}$$

Therefore,

$$\hat{\boldsymbol{b}} \sim N(\boldsymbol{b}, \sigma^2 (X^T X)^{-1}).$$

Since $\hat{\boldsymbol{e}} = \boldsymbol{Y} - X\hat{\boldsymbol{b}} = (I_n - X(X^TX)^{-1}X^T)\boldsymbol{Y}$, that is $\hat{\boldsymbol{e}}$ is a linear transformation of $\boldsymbol{Y}$, $\hat{\boldsymbol{e}}$ also follows a normal distribution with mean

$$\begin{aligned} E(\hat{\boldsymbol{e}}) &= E(\boldsymbol{Y} - X\hat{\boldsymbol{b}}) \\ &= E(\boldsymbol{Y} - X(X^TX)^{-1}X^T\boldsymbol{Y}) \\ &= E\left[(I_n - X(X^TX)^{-1}X^T)\boldsymbol{Y}\right] \\ &= (I_n - X(X^TX)^{-1}X^T)E(\boldsymbol{Y}) \\ &= (I_n - X(X^TX)^{-1}X^T)X\boldsymbol{b} \\ &= X\boldsymbol{b} - X(X^TX)^{-1}X^TX\boldsymbol{b} \\ &= \boldsymbol{0} \end{aligned}$$

and variance

$$\begin{aligned} Var(\hat{\boldsymbol{e}}) &= Cov\left[(I_n - X(X^TX)^{-1}X^T)\boldsymbol{Y},\ (I_n - X(X^TX)^{-1}X^T)\boldsymbol{Y}\right] \\ &= (I_n - X(X^TX)^{-1}X^T)Cov(\boldsymbol{Y},\ \boldsymbol{Y})(I_n - X(X^TX)^{-1}X^T) \\ &= \sigma^2(I_n - X(X^TX)^{-1}X^T) \end{aligned}$$

since $(I_n - X(X^TX)^{-1}X^T)$ is an idempotent matrix. Therefore,

$$\hat{\boldsymbol{e}} \sim N\left(\boldsymbol{0}, \sigma^2(I_n - X(X^TX)^{-1}X^T)\right).$$

For the distribution of $\hat{\sigma}^2$, note that

$$\hat{\sigma}^2 = \frac{\|\boldsymbol{Y} - X\hat{\boldsymbol{b}}\|^2}{(n-k-1)}$$

where

$$\begin{aligned} & \|\boldsymbol{Y} - X\hat{\boldsymbol{b}}\|^2 \\ = \ & (\boldsymbol{Y} - X\hat{\boldsymbol{b}})^T(\boldsymbol{Y} - X\hat{\boldsymbol{b}}) \\ = \ & \hat{\boldsymbol{e}}^T\hat{\boldsymbol{e}} \\ = \ & \left[(I_n - X(X^TX)^{-1}X^T)\boldsymbol{Y}\right]^T\left[(I_n - X(X^TX)^{-1}X^T)\boldsymbol{Y}\right] \\ = \ & \boldsymbol{Y}^T(I_n - X(X^TX)^{-1}X^T)\boldsymbol{Y} \qquad \text{(since } (I_n - X(X^TX)^{-1}X^T) \text{ is idempotent)} \\ = \ & (X\boldsymbol{b} + \boldsymbol{e})^T(I_n - X(X^TX)^{-1}X^T)(X\boldsymbol{b} + \boldsymbol{e}) \\ = \ & \boldsymbol{e}^T(I_n - X(X^TX)^{-1}X^T)\boldsymbol{e}. \end{aligned}$$

Note that

$$\begin{aligned} & trace(I_n - X(X^TX)^{-1}X^T) \\ = \ & trace(I_n) - trace(X(X^TX)^{-1}X^T) \\ = \ & n - trace\left((X^TX)^{-1}X^TX\right) \qquad \text{(since } trace(AB) = trace(BA)) \\ = \ & n - (k+1). \end{aligned}$$

Furthermore, $I_n - X(X^TX)^{-1}X^T$ is an idempotent matrix and so its rank is given by its trace. Therefore, there exists an $n \times n$ orthogonal matrix G such that

$$I_n - X(X^TX)^{-1}X^T = G^T \begin{pmatrix} I_{n-(k+1)} & 0 & 0 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 0 \end{pmatrix} G.$$

Let $D = (D_1, D_2, \ldots, D_n)^T = G\boldsymbol{e}/\sigma$ so that $G\boldsymbol{e} = \sigma D$. Since $\boldsymbol{e}/\sigma \sim N(\boldsymbol{0}, I_n)$ and G is orthogonal, $D \sim N(\boldsymbol{0}, I_n)$. Hence,

$$\begin{aligned} \|\boldsymbol{Y} - X\hat{\boldsymbol{b}}\|^2 &= \boldsymbol{e}^T(I_n - X(X^TX)^{-1}X^T)\boldsymbol{e} \\ &= \boldsymbol{e}^T G^T \begin{pmatrix} I_{n-(k+1)} & 0 & 0 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 0 \end{pmatrix} G\boldsymbol{e} \\ &= \sigma^2 D^T D \\ &= \sigma^2(D_1^2 + D_2^2 + \ldots + D_{n-k-1}^2) \\ &\sim \sigma^2 \chi^2_{n-k-1}. \end{aligned}$$

Therefore, the distribution of $\hat{\sigma}^2$ is given by

$$\hat{\sigma}^2 = \frac{\|\boldsymbol{Y} - X\hat{\boldsymbol{b}}\|^2}{n-k-1} \sim \frac{\sigma^2 \chi^2_{n-k-1}}{n-k-1}$$

1.2.3 Independence of $\hat{\boldsymbol{b}}$ and $\hat{\boldsymbol{e}}$ and independence of $\hat{\boldsymbol{b}}$ and $\hat{\sigma}^2$

Since $\hat{\boldsymbol{b}}$ and $\hat{\boldsymbol{e}}$ are normal random vectors and

$$\begin{aligned} Cov(\hat{\boldsymbol{b}}, \hat{\boldsymbol{e}}) &= Cov\left[(X^TX)^{-1}X^T\boldsymbol{Y}, (I_n - X(X^TX)^{-1}X^T)\boldsymbol{Y}\right] \\ &= \sigma^2(X^TX)^{-1}X^T(I_n - X(X^TX)^{-1}X^T) \\ &= 0, \end{aligned}$$

$\hat{\boldsymbol{b}}$ and $\hat{\boldsymbol{e}}$ are independent. Since $\hat{\sigma}^2 = \hat{\boldsymbol{e}}^T\hat{\boldsymbol{e}} \;/\; (n-k-1)$, $\hat{\boldsymbol{b}}$ and $\hat{\sigma}^2$ are also independent.

1.3 Uses of linear regression analysis

1.3.1 Hypothesis testing

Given a linear regression model, one can test whether some regression coefficients $\boldsymbol{b}$ satisfy certain constraints $H\boldsymbol{b} = \boldsymbol{h}$, where H is a given $r \times (k+1)$ matrix with full row rank $r \leq k+1$ and $\boldsymbol{h}$ is a given vector in $\Re^r$. For this, one tests

$$\begin{aligned} & H_0 : H\boldsymbol{b} = \boldsymbol{h} \\ \text{against} \quad & H_1 : H\boldsymbol{b} \neq \boldsymbol{h}. \end{aligned}$$

A size α test takes the form:

$$\text{Reject } \ H_0 \text{ if and only if } \ \frac{\left(\|\boldsymbol{Y} - X\hat{\boldsymbol{b}}_H\|^2 - \|\boldsymbol{Y} - X\hat{\boldsymbol{b}}\|^2\right)/r}{\|\boldsymbol{Y} - X\hat{\boldsymbol{b}}\|^2/(n-k-1)} > f^{\alpha}_{r,n-k-1}$$

where $\hat{\boldsymbol{b}}_H$ is the least squares estimate of $\boldsymbol{b}$, under the constraints $H\boldsymbol{b} = \boldsymbol{h}$ and $f^{\alpha}_{r,n-k-1}$ is the upper α point of an F distribution with degrees of freedom r and $n-k-1$. This test can also be derived as the likelihood ratio test.

One common hypothesis test is to assess whether a predictor variable of interest x_l add significantly to the prediction of the response y. In this case, the regression coefficient b_l is set to zero, H is set as a $1 \times (k+1)$ matrix with the $(1, l+1)$ entry equal to 1 and all the other entries equal to zero, and $\boldsymbol{h}$ is set to $\boldsymbol{0}$.

1.3.2 Confidence and prediction intervals

Confidence and prediction intervals provide useful ways of assessing the quality of prediction at a single point. When considering the mean response $\boldsymbol{x}^T\boldsymbol{b}$ of a model at a chosen $\boldsymbol{x} = (1, x_1, x_2, \ldots, x_k)^T$, a confidence interval can be constructed to provide useful information on where the mean response lies. Since $\boldsymbol{x}^T(\hat{\boldsymbol{b}} - \boldsymbol{b}) \sim N(\boldsymbol{0}, \sigma^2\boldsymbol{x}^T(X^TX)^{-1}\boldsymbol{x})$ and since $\hat{\boldsymbol{b}}$ is independent of $\hat{\sigma}^2$,

$$\frac{\boldsymbol{x}^T(\hat{\boldsymbol{b}} - \boldsymbol{b})}{\hat{\sigma}\sqrt{\boldsymbol{x}^T(X^TX)^{-1}\boldsymbol{x}}} \sim t_{n-k-1}$$

where t_{n-k-1} is a t distribution with $n-k-1$ degrees of freedom. Hence, a $1-\alpha$ confidence level for $\boldsymbol{x}^T\boldsymbol{b}$ has the form:

$$P\left\{\boldsymbol{x}^T\boldsymbol{b} \in \boldsymbol{x}^T\hat{\boldsymbol{b}} \pm t^{\alpha/2}_{n-k-1}\hat{\sigma}\sqrt{\boldsymbol{x}^T(X^TX)^{-1}\boldsymbol{x}}\right\} = 1 - \alpha$$

where $t^{\alpha/2}_{n-k-1}$ is the upper $\alpha/2$ point of a t distribution with $n-k-1$ degrees of freedom.

Moreover, when considering the value of future observation y_f at a chosen $\boldsymbol{x} = (1, x_1, x_2, \ldots, x_k)^T$, a prediction interval can be constructed to provide useful information on where y_f lies. In that case, $y_f = \boldsymbol{x}^T\boldsymbol{b} + e_f$ where e_f is the random error associated with y_f. Note that

$$\begin{aligned} y_f - \boldsymbol{x}^T\hat{\boldsymbol{b}} &= \boldsymbol{x}^T\boldsymbol{b} + e_f - \boldsymbol{x}^T\hat{\boldsymbol{b}} \\ &= e_f - \boldsymbol{x}^T(\hat{\boldsymbol{b}} - \boldsymbol{b}) \\ &\sim N(\boldsymbol{0}, \sigma^2 + \sigma^2\boldsymbol{x}^T(X^TX)^{-1}\boldsymbol{x}). \end{aligned}$$

since $e_f \sim N(0, \sigma^2)$ and is independent of $\hat{\boldsymbol{b}}$ and

$$\frac{y_f - \boldsymbol{x}^T\hat{\boldsymbol{b}}}{\hat{\sigma}\sqrt{1 + \boldsymbol{x}^T(X^TX)^{-1}\boldsymbol{x}}} \sim t_{n-k-1}.$$

Hence, a $1-\alpha$ prediction interval for y_f can be derived from

$$P\left\{y_f \in \boldsymbol{x}^T\hat{\boldsymbol{b}} \pm t^{\alpha/2}_{n-k-1}\hat{\sigma}\sqrt{1 + \boldsymbol{x}^T(X^TX)^{-1}\boldsymbol{x}}\right\} = 1 - \alpha.$$

1.4 Simultaneous confidence bands

When the entire range of values of the predictor variables is of interest, a simultaneous confidence band is used to make simultaneous confidence statements about the mean response $\boldsymbol{x}^T\boldsymbol{b}$ for all the possible values of $\boldsymbol{x}$ within the range of interest. A simultaneous confidence band provides useful information on where the true but unknown regression lies. A linear regression model $\boldsymbol{x}^T\boldsymbol{b}_0$ is a plausible candidate for the unknown regression $\boldsymbol{x}^T\boldsymbol{b}$ if and only if $\boldsymbol{x}^T\boldsymbol{b}_0$ is contained completely inside the confidence band

Simultaneous confidence bands can take various forms depending on the preferences and requirements of the user. For instance, if one wishes to have a confidence band which has simultaneous coverage probability of $1-\alpha$ and whose width is proportional to the standard error of the estimated regression function, then the Scheffé band can be used (see Scheffé, 1953). It has the form:

$$\boldsymbol{x}^T\boldsymbol{b} \in \boldsymbol{x}^T\hat{\boldsymbol{b}} \pm \hat{\sigma}\sqrt{(k+1)f^{\alpha}_{k+1,n-k-1}}\sqrt{\boldsymbol{x}^T(X^TX)^{-1}\boldsymbol{x}} \quad \forall \boldsymbol{x} \in \Re^{k+1}$$

where $f^{\alpha}_{k+1,n-k-1}$ is the upper α point of the F distribution with degrees of freedom $k+1$ and $n-k-1$.

Simultaneous confidence bands have now become a standard form of graphical illustration of results from statistical analysis. Although they are widely used nowadays, their construction is still a difficult problem for various regression models. Hence, the construction and comparison of exact two-sided confidence bands for various linear regression models are the main problems addressed in this thesis.

1.5 Outline of the thesis

The thesis consists of two main parts. The first part, which includes Chapter 2 and Chapter 3, covers the exact construction and optimality for two-sided simultaneous confidence bands in simple linear regression. The second part, which includes Chapter 4 and Chapter 5, covers the exact construction and optimality for two-sided simultaneous confidence bands in multiple linear regression. Specifically, in Chapter 2, the construction and comparison of some exact simultaneous confidence bands for simple linear regression are reviewed. Analytical and numerical methods are also carried out to show that D-optimal designs lead to the best confidence bands under a certain optimality criterion. In Chapter 3, two new families of simultaneous confidence bands are introduced and the best band within each family is identified numerically and compared to frequently quoted bands in the statistical literature. In Chapter 4, the construction and comparison of exact simultaneous two-sided confidence bands in multiple linear regression when the predictor variables are constrained in an ellipsoidal region are reviewed and a new family of simultaneous confidence bands is also introduced. A method to construct and compare of exact simultaneous two-sided confidence bands in multiple linear regression when the predictor variables are constrained in a rectangular region is proposed in Chapter 5. Finally, concluding remarks and possible future work are presented in Chapter 6.

Chapter 2

Exact simultaneous confidence bands in simple linear regression

2.1 The simple linear regression model

In simple linear regression, there is only one predictor variable in the regression model. The general model (1.1) is reduced to the response variable y being expressed as a function of the independent variable x and corresponding parameters b_0, b_1 and the error term e:

$$y = b_0 + b_1 x + e. \tag{2.1}$$

For a sample of n observations where the j^{th} observation is given by (y_j, x_j) for $j = 1, \ldots, n$, the j^{th} observation is assumed to satisfy the relationship:

$$y_j = b_0 + b_1 x_j + e_j.$$

The model in matrix form is the same as (1.2), that is $\boldsymbol{Y} = X\boldsymbol{b} + \boldsymbol{e}$. However, in simple linear regression, X denotes the design matrix whose i^{th} row is given by $(1, x_i)$ for $i = 1, \ldots, n$. So,

$$X = \begin{pmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_n \end{pmatrix}$$

and

$$(X^T X)^{-1} = \frac{1}{\sum_{j=1}^{n}(x_j - \bar{x})^2} \begin{pmatrix} \frac{\sum_{j=1}^{n} x_j^2}{n} & -\bar{x} \\ -\bar{x} & 1 \end{pmatrix}$$

where $\bar{x} = \frac{1}{n}\sum_{j=1}^{n} x_j$. The least squares estimates of $\boldsymbol{b} = (b_0, b_1)^T$ and σ are denoted by $\hat{\boldsymbol{b}} = (\hat{b}_0, \hat{b}_1)^T$ and $\hat{\sigma}$ respectively. In the simple linear regression case, $\hat{\boldsymbol{b}} \sim N_2(\boldsymbol{b}, \sigma^2(X^T X)^{-1})$ and $\frac{\hat{\sigma}}{\sigma} \sim \sqrt{\frac{\chi^2_\nu}{\nu}}$ where $\nu = n - k - 1 = n - 2$.

A $1-\alpha$ level confidence band $(l(x), u(x))$ for the regression line $b_0 + b_1 x$ over an interval $x \in (a, A)$ has the form

$$\inf_{-\infty < b_0, b_1 < \infty, \sigma > 0} P\{l(x) < b_0 + b_1 x < u(x) \quad \forall x \in (a, A)\} = 1 - \alpha \tag{2.2}$$

where $l(x)$ and $u(x)$ are given functions representing the lower and upper parts respectively of the band, and $-\infty \leq a < A \leq \infty$ and $\alpha \in (0,1)$ are given constants. The construction of simultaneous confidence bands for simple linear regression dates back to Working and Hotelling (1929) who obtained confidence bands for linear models when the variance is known. Since then, a considerable literature and work in the field has been made. For the case of unknown variance, Scheffé (1953) constructed hyperbolic confidence bands and Bowden and Graybill (1966) constructed straight line confidence bands. Wynn and Bloomfield (1971) and Uusipaikka (1983) provided exact confidence bands with width proportional to the standard error when the only predictor variable is restricted to an interval or union of intervals. Recent papers include Liu and Hayter(2007) who compare between confidence bands in simple linear regression and Liu, Lin and Piegorsch (2008) who provide methods for the construction of exact simultaneous confidence bands in simple linear regression.

In this chapter, we review the construction of the three most frequently quoted exact simultaneous confidence bands in the statistical literature, namely the two-segment band, the three-segment band and the hyperbolic band. More specifically, the two-sided confidence bands are considered, where the bands are symmetric about the estimated regression line $\hat{b}_0 + \hat{b}_1 x$ and $l(x)$ and $u(x)$ are given by

$$l(x) = \hat{b}_0 + \hat{b}_1 x - \hat{\sigma} H(x), \; u(x) = \hat{b}_0 + \hat{b}_1 x + \hat{\sigma} H(x)$$

where $\hat{\sigma} H(x) > 0$ is the half width of the band at x and $H(x)$ determines the shape of the band.

2.2 Preliminaries

Before focusing on the construction of each of the three bands, some preliminary results used in the construction of the bands are presented. Let U be the unique square root matrix of $(X^T X)^{-1}$ so that $(X^T X)^{-1} = U^2$. Using the results from Section 1.2.2 and applying them to simple linear regression, it is clear that

$$\hat{\boldsymbol{b}} \sim N_2(\boldsymbol{b}, \sigma^2 (X^T X)^{-1}), \qquad \frac{\hat{\sigma}}{\sigma} \sim \sqrt{\frac{\chi^2_{n-2}}{n-2}}$$

and $\hat{\boldsymbol{b}}$ and $\hat{\sigma}$ are independent. Hence, $\boldsymbol{N} = U^{-1}(\hat{\boldsymbol{b}} - \boldsymbol{b})/\sigma \sim N_2(\boldsymbol{0}, I_2)$ and $\boldsymbol{T} = \boldsymbol{N}/(\frac{\hat{\sigma}}{\sigma}) = U^{-1}(\hat{\boldsymbol{b}} - \boldsymbol{b})/\hat{\sigma}$ follows a bivariate t distribution (see Tong, 1990) whose probability density function is given by

$$f_{\boldsymbol{T}}(t_1, t_2) = \frac{1}{2\pi} \left[1 + \frac{1}{\nu}(t_1^2 + t_2^2) \right]^{-\frac{(\nu+2)}{2}}, \quad (t_1, t_2) \in \Re^2.$$

Moreover, denote the polar coordinates of $\boldsymbol{N} = (N_1, N_2)^T$ and $\boldsymbol{T} = (t_1, t_2)^T$ by $(R_{\boldsymbol{N}}, \theta_{\boldsymbol{N}})$ and $(R_{\boldsymbol{T}}, \theta_{\boldsymbol{T}})$ respectively. For $R_{\boldsymbol{N}} \geq 0$ and $\theta_{\boldsymbol{N}} \in [0, 2\pi)$,

$$N_1 = R_{\boldsymbol{N}} \cos \theta_{\boldsymbol{N}},$$

$$N_2 = R_{\boldsymbol{N}} \sin \theta_{\boldsymbol{N}}.$$

Then, $R_{\boldsymbol{N}} = \sqrt{N_1^2 + N_2^2} \sim \sqrt{\chi_2^2}$, $\theta_{\boldsymbol{N}}$ has a uniform distribution on the interval $[0, 2\pi)$, and $R_{\boldsymbol{N}}$ and $\theta_{\boldsymbol{N}}$ are independent random variables. Therefore,

$$R_{\boldsymbol{T}} = R_{\boldsymbol{N}}/(\frac{\hat{\sigma}}{\sigma}) \ \sim \frac{\sqrt{\chi_2^2}}{\sqrt{\chi_{n-2}^2/(n-2)}} = \sqrt{2F_{2,n-2}}$$

where F_{k_1,k_2} denotes an F random variable with degrees of freedom k_1 and k_2. Furthermore, $\theta_{\boldsymbol{T}}$ has a uniform distribution on the interval $[0, 2\pi)$ and $R_{\boldsymbol{T}}$ and $\theta_{\boldsymbol{T}}$ are independent random variables. The cumulative distribution function of $R_{\boldsymbol{T}}$ is given by

$$\begin{aligned} F_{R_{\boldsymbol{T}}}(x) &= P\{R_{\boldsymbol{T}} < x\} \\ &= P\{\sqrt{2F_{2,n-2}} < x\} \\ &= P\{F_{2,n-2} < \frac{x^2}{2}\} \\ &= \int_0^{\frac{x^2}{2}} \left(1 + \frac{2u}{\nu}\right)^{-\frac{\nu+2}{2}} du \end{aligned}$$

since the pdf of the random variable $F_{2,n-2}$ is given by $f_{2,\nu}(u) = \left(1 + \frac{2u}{\nu}\right)^{-\frac{\nu+2}{2}}$, where $\nu = n - 2$. Therefore,

$$F_{R_{\boldsymbol{T}}}(x) = 1 - \left(1 + \frac{x^2}{\nu}\right)^{-\frac{\nu}{2}}. \tag{2.3}$$

For a given vector $\boldsymbol{v} \in \Re^2$ and constant $r > 0$, the set

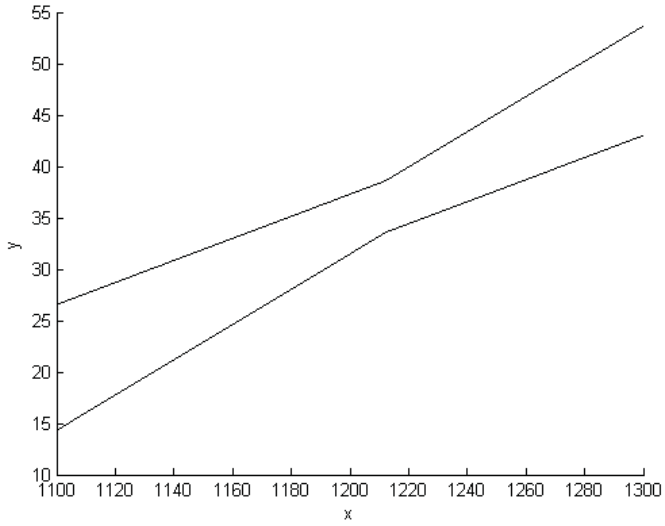
$$\left\{\boldsymbol{T} : \boldsymbol{v}^T\boldsymbol{T}/\|\boldsymbol{v}\| < c\right\} \subset R^2$$

is made up of all the points that are on the same side as the origin of the straight line $\boldsymbol{v}^T\boldsymbol{T}/\|\boldsymbol{v}\| = c$, where $\boldsymbol{v}^T\boldsymbol{T}/\|\boldsymbol{v}\| = c$ is perpendicular to the vector $\boldsymbol{v}$ and c distance away, in the direction of $\boldsymbol{v}$, from the origin. Hence, the set $\left\{\boldsymbol{T} : \left|\boldsymbol{v}^T\boldsymbol{T}\right|/\|\boldsymbol{v}\| < c\right\} \subset R^2$, which can be expressed as $\left\{\boldsymbol{T} : \boldsymbol{v}^T\boldsymbol{T}/\|\boldsymbol{v}\| < c\right\} \cap \left\{\boldsymbol{T} : (-\boldsymbol{v})^T\boldsymbol{T}/\|(-\boldsymbol{v})\| < c\right\}$, is the stripe bounded by the parallel lines $\boldsymbol{v}^T\boldsymbol{T}/\|\boldsymbol{v}\| = c$ and $\boldsymbol{v}^T\boldsymbol{T}/\|\boldsymbol{v}\| = -c$. Finally, define a function $v(c,d)$ as

$$\begin{aligned} v(c,d) &= Var\left\{(c,d)\hat{\boldsymbol{b}}\right\}/\sigma^2 \\ &= \begin{pmatrix} c & d \end{pmatrix} (X^TX)^{-1} \begin{pmatrix} c \\ d \end{pmatrix} \\ &= \left\{U \begin{pmatrix} c \\ d \end{pmatrix}\right\}^T \left\{U \begin{pmatrix} c \\ d \end{pmatrix}\right\} \\ &= \left\|U \begin{pmatrix} c \\ d \end{pmatrix}\right\|^2 \end{aligned}$$

which will be used in deriving the form and confidence level of the bands.

Figure 2.1: Two-sided two-segment band



2.3 Some exact simultaneous confidence bands

2.3.1 Two-sided two-segment bands

A two-sided two-segment band has

$$H_{2,2}(x) = c_{2,2,1}\sqrt{v(1,\bar{x})} + c_{2,2,2}|x-\bar{x}|\sqrt{v(0,1)}, \quad x \in (-\infty,\infty) \tag{2.4}$$

where the critical constants $c_{2,2,1}$ and $c_{2,2,2}$ are chosen so that the confidence level of the band is equal to $1-\alpha$. The form of the two-sided two-segment band is illustrated in Figure 2.1. The band satisfies the following probability:

$$P\{b_0 + b_1 x \in \hat{b_0} + \hat{b_1} x \pm \hat{\sigma} H_{2,2}(x) \ \ \forall x \in (-\infty,\infty)\} = 1-\alpha.$$

The probability can also be expressed as

$$P\left\{\sup_{x\in(-\infty,\infty)} \frac{|(1,x)(\hat{\boldsymbol{b}}-\boldsymbol{b})/\hat{\sigma}|}{H_{2,2}(x)} < 1\right\} = 1-\alpha.$$

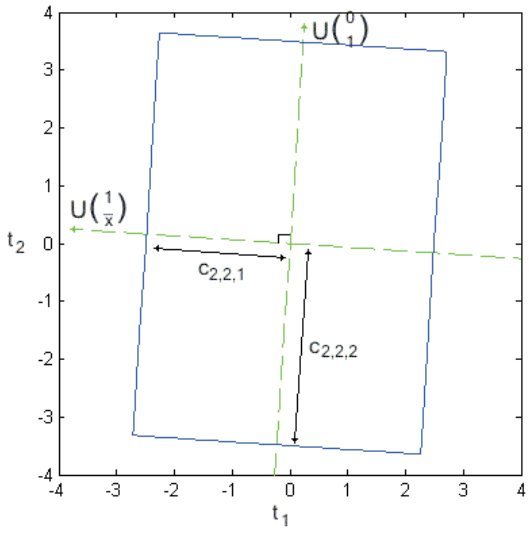
Note that

$$\frac{\partial}{\partial x}\left\{\left[\frac{(1,x)(\hat{\boldsymbol{b}}-\boldsymbol{b})/\hat{\sigma}}{H_{2,2}(x)}\right]\right\}$$

has a fixed sign, either positive or negative, over $x < \bar{x}$ and $x > \bar{x}$. The supreme is therefore attained at either $x = \bar{x}$ or limits $x \to -\infty$ or $x \to \infty$. So, the confidence level can be further expressed as

$$P\left\{\sup_{x=-\infty \text{ or } \bar{x} \text{ or } \infty} \frac{|(1,x)(\hat{\boldsymbol{b}}-\boldsymbol{b})/\hat{\sigma}|}{H_{2,2}(x)} < 1\right\} = P\{\boldsymbol{T} \in R_2\}$$

Figure 2.2: The region R_2



where $R_2 = R_2(-\infty) \cap R_2(\bar{x}) \cap R_2(\infty)$ with

$$\begin{aligned} R_2(\bar{x}) &= \left\{ \boldsymbol{T} : \frac{|(1,\bar{x})(\hat{\boldsymbol{b}} - \boldsymbol{b})/\hat{\sigma}|}{H_{2,2}(\bar{x})} < 1 \right\} \\ &= \left\{ \boldsymbol{T} : \left| \left\{ U \begin{pmatrix} 1 \\ \bar{x} \end{pmatrix} \right\}^T \boldsymbol{T} \right| \; / \; c_{2,2,1}\sqrt{v(1,\bar{x})} < 1 \right\} \\ &= \left\{ \boldsymbol{T} : \left| \left\{ U \begin{pmatrix} 1 \\ \bar{x} \end{pmatrix} \right\}^T \boldsymbol{T} \right| \; / \; \left\| U \begin{pmatrix} 1 \\ \bar{x} \end{pmatrix} \right\| < c_{2,2,1} \right\} \end{aligned}$$

and

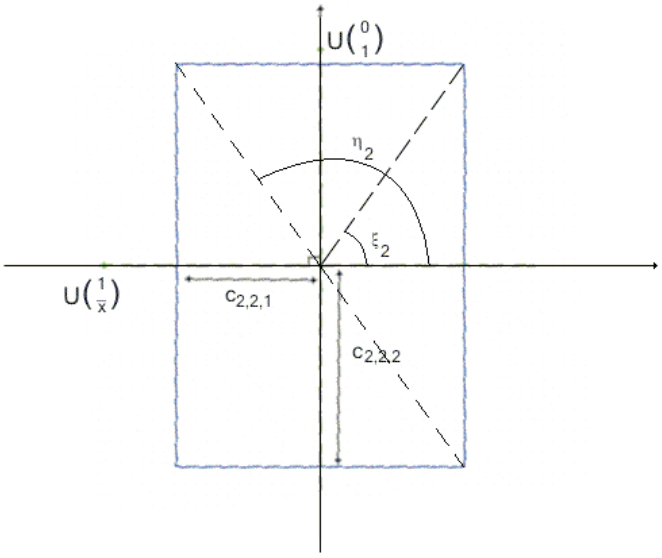
$$\begin{aligned} R_2(-\infty) = R_2(\infty) &= \left\{ \boldsymbol{T} : \lim_{x \to \infty} \frac{|(1,x)(\hat{\boldsymbol{b}} - \boldsymbol{b})/\hat{\sigma}|}{H_{2,2}(x)} < 1 \right\} \\ &= \left\{ \boldsymbol{T} : \left| \left\{ U \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\}^T \boldsymbol{T} \right| \; / \; \left\| U \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\| < c_{2,2,2} \right\}. \end{aligned}$$

Using the results from Section 2.2, R_2 is hence found to be the region given by a parallelogram whose sides are given by the lines which are $c_{2,2,2}$ and $c_{2,2,1}$ distance away from the origin and perpendicular to the vectors $U(0,1)^T$ and $U(1,\bar{x})^T$ respectively. R_2 is illustrated in Figure 2.2. The angle formed by the vectors $U(0,1)^T$ and $U(1,\bar{x})^T$ is $\pi/2$ since the cosine of that angle is given by

$$\frac{\left\{ U \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\}^T \left\{ U \begin{pmatrix} 1 \\ \bar{x} \end{pmatrix} \right\}}{\left\| U \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\| \left\| U \begin{pmatrix} 1 \\ \bar{x} \end{pmatrix} \right\|} = \frac{\begin{pmatrix} 0 & 1 \end{pmatrix} (X^T X)^{-1} \begin{pmatrix} 1 \\ \bar{x} \end{pmatrix}}{\left\| U \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\| \left\| U \begin{pmatrix} 1 \\ \bar{x} \end{pmatrix} \right\|} = 0.$$

Therefore, R_2 is given by a rectangular region. The confidence level of the two-sided two-segment band is given by the probability of $\boldsymbol{T}$ in R_2. Let R_2^* be the region that

Figure 2.3: The region R_2^*



is resulted from rotating R_2 around the origin to the position so that $U(1, \bar{x})^T$ is in the direction of the t_1-axis and $U(0, 1)^T$ is in the direction of the t_2-axis, as shown in Figure 2.3. Due to the rotational invariance of the probability distribution of $\boldsymbol{T}$, the probability of $\boldsymbol{T}$ in R_2^* is equal to the probability of $\boldsymbol{T}$ in R_2. Furthermore, the probability of $\boldsymbol{T}$ in R_2^* is equal to twice the probability of $\boldsymbol{T}$ in the top-right half of R_2^*, which can be expressed as

$$\{\boldsymbol{T} : \theta_{\boldsymbol{T}} \in [-(\pi - \eta_2), \xi_2] \ , \ R_{\boldsymbol{T}} \cos\theta_{\boldsymbol{T}} \leq c_{2,2,1}\} \cup \{\boldsymbol{T} : \theta_{\boldsymbol{T}} \in [\xi_2, \eta_2] \ , \ R_{\boldsymbol{T}} \cos(\theta_{\boldsymbol{T}} - \pi/2) \leq c_{2,2,2}\}$$

where the angles ξ_2 and η_2 are depicted in Figure 2.3 and given by

$$\xi_2 = \sin^{-1}\left(\frac{c_{2,2,2}}{\sqrt{c_{2,2,2}^2 + c_{2,2,1}^2}}\right)$$

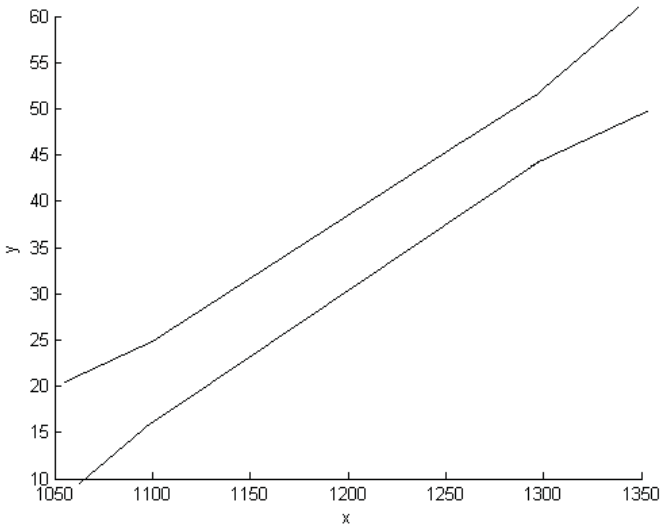
and

$$\eta_2 = \cos^{-1}\left(\frac{-c_{2,2,1}}{\sqrt{c_{2,2,2}^2 + c_{2,2,1}^2}}\right).$$

Hence, the probability of $\boldsymbol{T}$ in R_2 is given by

$$\begin{aligned}
& P\{\boldsymbol{T} \in R_2\} \\
= \ & 2P\left\{\boldsymbol{T} : \theta_{\boldsymbol{T}} \in [-(\pi - \eta_2), \xi_2] \ , \ R_{\boldsymbol{T}} \cos\theta_{\boldsymbol{T}} \leq c_{2,2,1}\right\} \\
+ \ & 2P\left\{\boldsymbol{T} : \theta_{\boldsymbol{T}} \in [\xi_2, \eta_2] \ , \ R_{\boldsymbol{T}} \cos(\theta_{\boldsymbol{T}} - \pi/2) \leq c_{2,2,2}\right\} \\
= \ & 2\int_{-(\pi-\eta_2)}^{\xi_2} \frac{1}{2\pi} P\{R_{\boldsymbol{T}} \cos\theta \leq c_{2,2,1}\}\, d\theta \\
+ \ & 2\int_{\xi_2}^{\eta_2} \frac{1}{2\pi} P\{R_{\boldsymbol{T}} \cos\left(\theta - \frac{\pi}{2}\right) \leq c_{2,2,2}\}\, d\theta \qquad (2.5) \\
= \ & \frac{1}{\pi}\int_{-(\pi-\eta_2)}^{\xi_2} F_{R_{\boldsymbol{T}}}\left(\frac{c_{2,2,1}}{\cos\theta}\right) d\theta + \frac{1}{\pi}\int_{\xi_2 - \frac{\pi}{2}}^{\eta_2 - \frac{\pi}{2}} F_{R_{\boldsymbol{T}}}\left(\frac{c_{2,2,2}}{\cos\theta}\right) d\theta \qquad (2.6)
\end{aligned}$$

Figure 2.4: Two-sided three-segment band



where the function $F_{R_{\boldsymbol{T}}}(x)$ is defined by expression (2.3), equality (2.5) follows directly from the uniform distribution of $\theta_{\boldsymbol{T}}$ and equality (2.6) follows directly from the cumulative distribution function of $R_{\boldsymbol{T}}$.

2.3.2 Two-sided three-segment bands

A two-sided three-segment band has

$$H_{3,2}(x) = \frac{1}{A-a}\{(x-a)c_{3,2,1}\sqrt{v(1,A)} + (A-x)c_{3,2,2}\sqrt{v(1,a)}\}, \quad x \in (a,A) \qquad (2.7)$$

where the critical constants $c_{3,2,1}$ and $c_{3,2,2}$ are chosen so that the confidence level of the band is equal to $1-\alpha$. The form of the two-sided three-segment band is illustrated in Figure 2.4. For x outside of (a,A), the band is formed of straight lines corresponding to the diagonal extensions of the band within (a,A) and thus, the upper and lower parts of the band each consists of three line segments. The band satisfies

$$P\{b_0 + b_1 x \in \hat{b_0} + \hat{b_1}x \pm \hat{\sigma}H_{3,2}(x)\ \forall x \in (a,A)\} = 1-\alpha.$$

The probability on the left side of the equality can be expressed as

$$P\left\{\sup_{x\in(a,A)} \frac{|(1,x)(\hat{\boldsymbol{b}}-\boldsymbol{b})/\hat{\sigma}|}{H_{3,2}(x)} < 1\right\}.$$

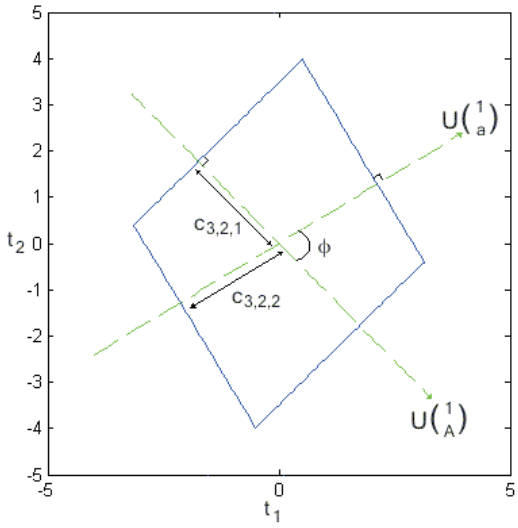
As in section 2.3.1,

$$\frac{\partial}{\partial x}\left\{\left[\frac{(1,x)(\hat{\boldsymbol{b}}-\boldsymbol{b})/\hat{\sigma}}{H_{3,2}(x)}\right]\right\}$$

has a fixed sign, either positive or negative, over $x \in (a,A)$. The supreme is therefore attained at either $x=a$ or $x=A$. So, the confidence level can be further expressed as

$$P\left\{\sup_{x=a\ or\ A} \frac{|(1,x)(\hat{\boldsymbol{b}}-\boldsymbol{b})/\hat{\sigma}|}{H_{3,2}(x)} < 1\right\} = P\{\boldsymbol{T} \in R_3\}$$

Figure 2.5: The region R_3



where $R_3 = R_3(a) \cap R_3(A)$ with

$$\begin{aligned} R_3(a) &= \left\{ \boldsymbol{T} : \frac{|(1,a)(\hat{\boldsymbol{b}} - \boldsymbol{b})/\hat{\sigma}|}{H_{3,2}(a)} < 1 \right\} \\ &= \left\{ \boldsymbol{T} : \left| \left\{ U \begin{pmatrix} 1 \\ a \end{pmatrix} \right\}^T \boldsymbol{T} \right| \; / \; c_{3,2,2}\sqrt{v(1,a)} < 1 \right\} \\ &= \left\{ \boldsymbol{T} : \left| \left\{ U \begin{pmatrix} 1 \\ a \end{pmatrix} \right\}^T \boldsymbol{T} \right| \; / \; \left\| U \begin{pmatrix} 1 \\ a \end{pmatrix} \right\| < c_{3,2,2} \right\}. \end{aligned}$$

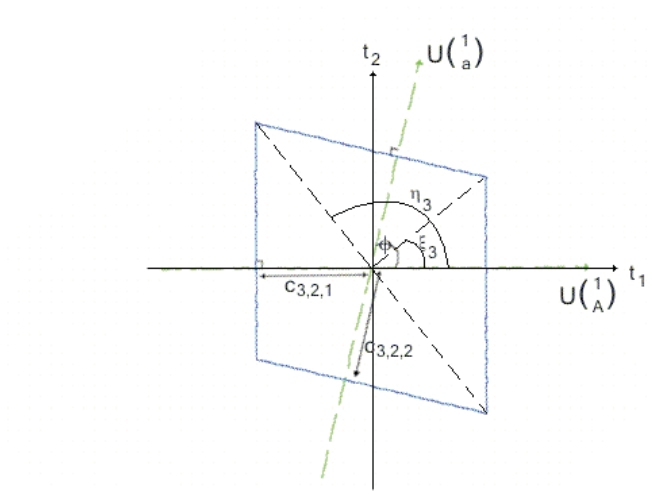
Similarly,

$$\begin{aligned} R_3(A) &= \left\{ \boldsymbol{T} : \left| \left\{ U \begin{pmatrix} 1 \\ A \end{pmatrix} \right\}^T \boldsymbol{T} \right| \; / \; c_{3,2,1}\sqrt{v(1,A)} < 1 \right\} \\ &= \left\{ \boldsymbol{T} : \left| \left\{ U \begin{pmatrix} 1 \\ A \end{pmatrix} \right\}^T \boldsymbol{T} \right| \; / \; \left\| U \begin{pmatrix} 1 \\ A \end{pmatrix} \right\| < c_{3,2,1} \right\}. \end{aligned}$$

The region R_3 is given by a parallelogram whose sides are given by the lines which are $c_{3,2,1}$ and $c_{3,2,2}$ distance away from the origin and perpendicular to the vectors $U(1,a)^T$ and $U(1,A)^T$ respectively. The region R_3 is illustrated in Figure 2.5. The angle ϕ is formed by the vectors $U(1,a)^T$ and $U(1,A)^T$ and can be calculated from

$$\begin{aligned} \cos\phi &= \frac{\left\{ U \begin{pmatrix} 1 \\ a \end{pmatrix} \right\}^T \left\{ U \begin{pmatrix} 1 \\ A \end{pmatrix} \right\}}{\left\| U \begin{pmatrix} 1 \\ a \end{pmatrix} \right\| \left\| U \begin{pmatrix} 1 \\ A \end{pmatrix} \right\|} \qquad (2.8) \\ &= \frac{\begin{pmatrix} 1 & a \end{pmatrix} (X^T X)^{-1} \begin{pmatrix} 1 \\ A \end{pmatrix}}{\sqrt{v(1,a)v(1,A)}}. \end{aligned}$$

Figure 2.6: The region R_3^*



Note that $\cos\phi$ is also the correlation coefficient between $\hat{b_0} + \hat{b_1}a$ and $\hat{b_0} + \hat{b_1}A$. The confidence level of the two-sided three-segment band is given by the probability of $\boldsymbol{T}$ in R_3.

Let R_3^* be the region that is resulted from rotating R_3 around the origin to the position so that $U(1, A)^T$ is in the direction of the t_1-axis, as shown in Figure 2.6. Due to the rotational invariance of the probability distribution of $\boldsymbol{T}$, the probability of $\boldsymbol{T}$ in R_3 is equal to the probability of $\boldsymbol{T}$ in R_3^*, which is further equal to twice the probability of $\boldsymbol{T}$ in the top-right half of R_3^*, which can be expressed as

$$\{\boldsymbol{T} : \theta_{\boldsymbol{T}} \in [-(\pi - \eta_3), \xi_3] \ , \ R_{\boldsymbol{T}} \cos\theta_{\boldsymbol{T}} \leq c_{3,2,1}\} \cup \{\boldsymbol{T} : \theta_{\boldsymbol{T}} \in [\xi_3, \eta_3] \ , \ R_{\boldsymbol{T}} \cos(\theta_{\boldsymbol{T}} - \phi) \leq c_{3,2,2}\}$$

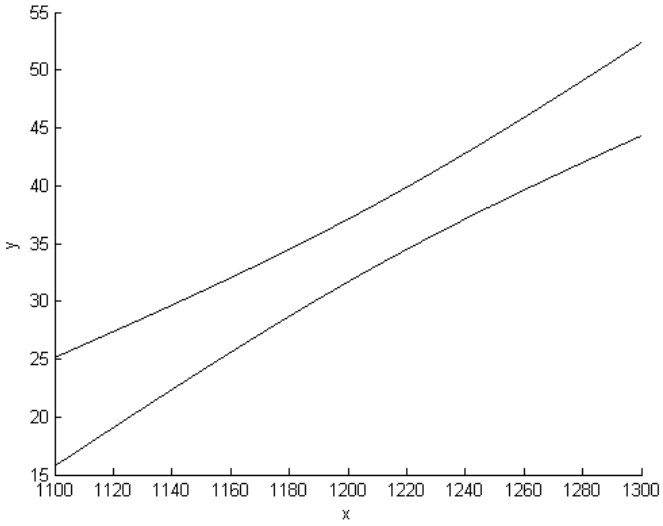
where angles ξ_3 and η_3 can be derived using trigonometric calculations to be:

$$\xi_3 = \sin^{-1}\left(\frac{c_{3,2,2} - c_{3,2,1}\cos\phi}{\sqrt{c_{3,2,2}^2 + c_{3,2,1}^2 - 2c_{3,2,2}c_{3,2,1}\cos\phi}}\right)$$

and

$$\eta_3 = \cos^{-1}\left(\frac{-c_{3,2,1}\sin\phi}{\sqrt{c_{3,2,2}^2 + c_{3,2,1}^2 + 2c_{3,2,2}c_{3,2,1}\cos\phi}}\right).$$

Figure 2.7: Two-sided hyperbolic band



Hence, the probability of $\boldsymbol{T}$ in R_3 is equal to

$$\begin{aligned}
& P\{\boldsymbol{T} \in R_3\} \\
= \quad & 2P\left\{\boldsymbol{T} : \theta_{\boldsymbol{T}} \in [-(\pi-\eta_3), \xi_3]\ ,\ R_{\boldsymbol{T}} \cos\theta_{\boldsymbol{T}} \le c_{3,2,1}\right\} \\
+ \quad & 2P\left\{\boldsymbol{T} : \theta_{\boldsymbol{T}} \in [\xi_3, \eta_3]\ ,\ R_{\boldsymbol{T}} \cos\left(\theta_{\boldsymbol{T}} - \phi\right) \le c_{3,2,2}\right\} \\
= \quad & 2\int_{-(\pi-\eta_3)}^{\xi_3} \frac{1}{2\pi} P\{R_{\boldsymbol{T}} \cos\theta \le c_{3,2,1}\}\, d\theta \\
+ \quad & 2\int_{\xi_3}^{\eta_3} \frac{1}{2\pi} P\{R_{\boldsymbol{T}} \cos\left(\theta-\phi\right) \le c_{3,2,2}\}\, d\theta \\
= \quad & \frac{1}{\pi}\int_{-(\pi-\eta_3)}^{\xi_3} F_{R_{\boldsymbol{T}}}\left(\frac{c_{3,2,1}}{\cos\theta}\right) d\theta + \frac{1}{\pi}\int_{\xi_3-\phi}^{\eta_3-\phi} F_{R_{\boldsymbol{T}}}\left(\frac{c_{3,2,2}}{\cos\theta}\right) d\theta
\end{aligned} \tag{2.9}$$

where the function $F_{R_{\boldsymbol{T}}}(x)$ is defined by expression (2.3). Thus, the confidence level of a two-sided three-segment band is given by expression (2.9).

2.3.3 Two-sided hyperbolic bands

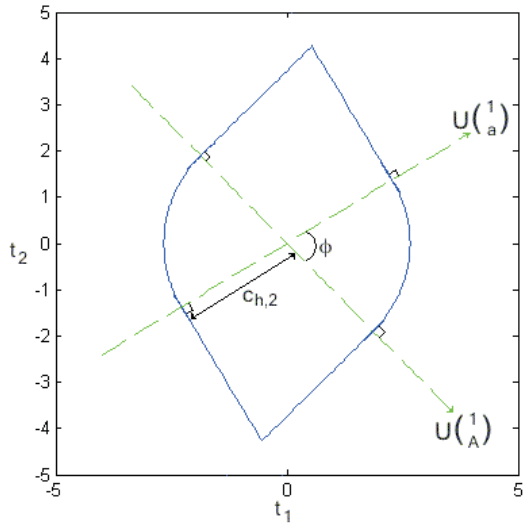
A two-sided hyperbolic band has

$$H_{h,2}(x) = c_{h,2}\sqrt{v(1,x)}, \quad x \in (a, A) \tag{2.10}$$

where the critical constant $c_{h,2}$ is chosen so that the confidence level of the band is equal to $1-\alpha$. The form of the two-sided hyperbolic band is illustrated in Figure 2.7. The band satisfies

$$P\{b_0 + b_1 x \in \hat{b_0} + \hat{b_1} x \pm c_{h,2}\hat{\sigma}\sqrt{v(1,x)}\ \forall x \in (a, A)\} = 1-\alpha$$

Figure 2.8: The region R_h



$$
\begin{aligned}
&= P\left\{ \sup_{x\in(a,A)} \frac{|\begin{pmatrix} 1 & x \end{pmatrix}(\hat{\boldsymbol{b}}-\boldsymbol{b})/\hat{\sigma}|}{\sqrt{v(1,x)}} < c_{h,2} \right\} \\
&= P\left\{ \sup_{x\in(a,A)} \left| \left\{ U \begin{pmatrix} 1 \\ x \end{pmatrix} \right\}^T \boldsymbol{T} \right| \Big/ \left\| U \begin{pmatrix} 1 \\ x \end{pmatrix} \right\| < c_{h,2} \right\} \\
&= P\{\boldsymbol{T} \in R_h\}
\end{aligned}
$$

where $R_h = \cap_{x\in(a,A)} R_h(x)$ with

$$
R_h(x) = \left\{ \boldsymbol{T} : \left| \left\{ U \begin{pmatrix} 1 \\ x \end{pmatrix} \right\}^T \boldsymbol{T} \right| \Big/ \left\| U \begin{pmatrix} 1 \\ x \end{pmatrix} \right\| < c_{h,2} \right\}.
$$

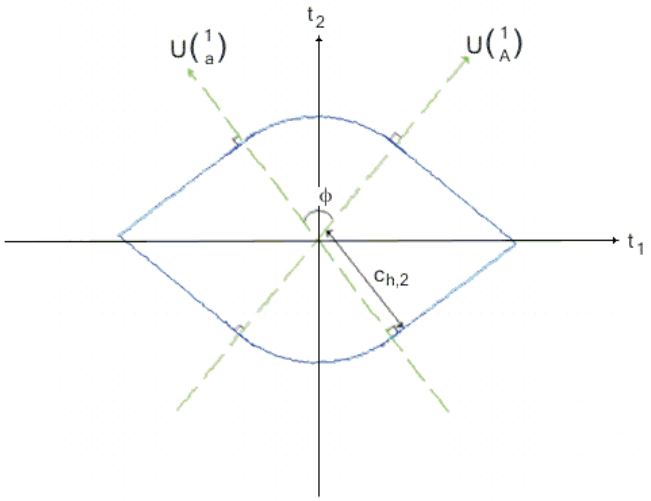
The region $R_h(x)$ is given by a strip bounded by the 2 lines that are $c_{h,2}$ distance away from the origin and perpendicular to the vector $U(1,x)^T$. Therefore, R_h is the region given by a spindle region whose angle at the vertices is ϕ, as depicted in Figure 2.8, where ϕ is also the angle between $U(1,a)^T$ and $U(1,A)^T$ and is calculated as in Section 2.3.2 for the two-sided three-segment band. The confidence level of the two-sided hyperbolic band is given by the probability of $\boldsymbol{T}$ in R_h. Let R_h^* be the region that is resulted from rotating R_h around the origin to the position so that the angle ϕ between $U(1,a)^T$ and $U(1,A)^T$ is divided into two equal halves by the t_2-axis, as shown in Figure 2.9. Due to the rotational invariance of the probability distribution of $\boldsymbol{T}$, the probability of $\boldsymbol{T}$ in R_h is equal to the probability of $\boldsymbol{T}$ in R_h^*. The region R_h^* is divided by the axes into four equal quarters. The top-right quarter of R_h^*, for instance, can be partitioned into two parts, the fan given by

$$
\left\{ \boldsymbol{T} : \theta_{\boldsymbol{T}} \in \left[\frac{\pi-\phi}{2}, \frac{\pi}{2} \right] \ , \ \|\boldsymbol{T}\| \le c_{h,2} \right\}
$$

and the right-angle triangle given by

$$
\left\{ \boldsymbol{T} : \theta_{\boldsymbol{T}} \in \left[0, \frac{\pi-\phi}{2} \right] \ , \ 0 \le \left(\cos\left(\frac{\pi-\phi}{2} \right), \sin\left(\frac{\pi-\phi}{2} \right) \right) \boldsymbol{T} \le c_{h,2} \right\}.
$$

Figure 2.9: The region R_h^*



Therefore, the probability of $\boldsymbol{T}$ in R_h is equal to four times the sum of the probabilities of $\boldsymbol{T}$ in the two regions mentioned above and given by

$$\begin{aligned}
& P\{\boldsymbol{T} \in R_h\} \\
= \quad & 4P\left\{\theta_{\boldsymbol{T}} \in \left[0, \frac{\pi-\phi}{2}\right] \ , \ 0 \le \left(\cos\left(\frac{\pi-\phi}{2}\right), \sin\left(\frac{\pi-\phi}{2}\right)\right)\boldsymbol{T} \le c_{h,2}\right\} \\
+ \quad & 4P\left\{\theta_{\boldsymbol{T}} \in \left[\frac{\pi-\phi}{2}, \frac{\pi}{2}\right] \ , \ \|\boldsymbol{T}\| \le c_{h,2}\right\} \\
= \quad & 4\left[\int_0^{\frac{\pi-\phi}{2}} \frac{1}{2\pi} P\left\{0 \le R_{\boldsymbol{T}} \le \frac{c_{h,2}}{\cos\left(\frac{\pi-\phi}{2} - \theta\right)}\right\} d\theta \ + \ \frac{\left(\frac{\phi}{2}\right)}{2\pi} P\{R_{\boldsymbol{T}} \le c_{h,2}\}\, d\theta\right] \\
= \quad & 4\left[\int_0^{\frac{\pi-\phi}{2}} \frac{1}{2\pi}\left\{F_{R_{\boldsymbol{T}}}\left(\frac{c_{h,2}}{\sin(\theta + \frac{\phi}{2})}\right) - F_{R_{\boldsymbol{T}}}(0)\right\} d\theta \ + \ \frac{\left(\frac{\phi}{2}\right)}{2\pi} F_{R_{\boldsymbol{T}}}(c_{h,2})\right]
\end{aligned}$$

where the function $F_{R_{\boldsymbol{T}}}(x)$ is defined by expression (2.3), so that the confidence level of a two-sided hyperbolic band is given by

$$1 - \frac{\phi}{\pi}\left(1 + \frac{c_{h,2}^2}{\nu}\right)^{-\frac{\nu}{2}} - \frac{2}{\pi}\int_0^{\frac{\pi-\phi}{2}} \left[1 + \frac{c_{h,2}^2}{\nu \sin^2(\theta + \frac{\phi}{2})}\right]^{-\frac{\nu}{2}} d\theta. \tag{2.11}$$

There are more than one way to derive the confidence level of each of the types of confidence bands discussed in this chapter. Derivation of confidence level for one-sided confidence bands can be achieved using similar methods. Moreover, the expressions (2.6), (2.9) and (2.11) for the confidence levels of two-sided confidence bands for simple linear regression in this chapter involve computation no harder than one-dimensional integration. Next, we discuss the optimality criteria for the two-segment, three-segment and hyperbolic bands.

2.4 Optimality criteria for simultaneous confidence bands

An optimality criterion is a single expression that summarizes how good the entity being assessed is. The entity under a criterion is said to be optimal when it is maximized or

minimized depending on the requirements of the user. Literature on optimality criteria for experimental design is widely available and Atkinson, Donev and Tobias (2007) have provided a useful account of some of the most important ones. However, for simultaneous confidence bands, there exist fewer optimality criteria and two of them are outlined in this section. Then, the design leading to the best confidence bands under one of these criteria is derived analytically or numerically, for the two-sided two-segment, three-segment and hyperbolic bands.

2.4.1 Average width criterion

The idea of average width of a band as criterion was introduced by Gafarian (1964). It was then formalized by Naiman (1984), who defines simultaneous confidence bands for linear regression functions as μ-optimal among a family of confidence bands, if they minimize the average width of the bands with respect to the probability measure μ over the range of interest, among all the confidence bands in the family with equal coverage probability. For a two-sided simultaneous confidence band for an unknown regression function $f(x)$ over the region of interest with coverage probability $1-\alpha$, the points $x_1, \ldots, x_n$ where the band will be used are random vectors which are identically and independently distributed according to the probability measure μ. Therefore, the accuracy of the resulting confidence band is proportional to $\sum_{i=1}^{n} \frac{H(x)}{n}$, which converges almost surely to $\int_x H(x)\mu\, dx$ by the strong law of large numbers, provided the integral is finite. Thus, a μ-optimal simultaneous confidence band is the one with optimal average width accuracy. Intuitively, an optimal simultaneous confidence band under the average width criterion bounds the regression function over the range of interest as tightly as possible. This optimality criterion is used in Chapter 5 to compare between two types of confidence bands for a particular dataset.

Most work on confidence bands use the average width as optimality criterion. However, Liu and Hayter (2007) pointed out two flaws in the criterion. When comparing between confidence bands, they found that whichever band with the smaller critical constant will be deemed as the better band under the average width criterion. Moreover, the range of interest is a crucial factor when comparing bands under the criterion. For three-segment bands for instance, a three-segment band can be deemed better than another three-segment band depending on the ranges of interest that are used, although the pairs of simultaneous confidence intervals underlying the two three-segment bands are fixed.

2.4.2 Minimum area confidence set criterion

The Minimum Area Confidence Set (MACS) criterion for simultaneous confidence bands was introduced by Liu and Hayter (2007), who have defined a simultaneous confidence band for simple linear regression as optimal if the corresponding set for the linear parameters of regression model has the smallest area, among all confidence sets corresponding to simultaneous confidence bands with equal coverage probability $1-\alpha$. For multiple linear regression, they use the analogous Minimum Volume Confidence Set (MVCS) criterion. Each $1-\alpha$ level confidence band correspond to a $1-\alpha$ level confidence set for $\boldsymbol{b}$ (see e.g.,

Khorasani and Milliken, 1979 and Piegorsch, 1987). Intuitively, in simple linear regression for instance, each point (b_0, b_1) within a $1-\alpha$ level confidence set correspond to a line $b_0 + b_1 x$ lying completely within the $1-\alpha$ level confidence band. The smaller the area of the confidence set the fewer the candidates for the true and unknown regression line there are in the corresponding confidence band, and thus, the better the band is.

The MACS criterion is related to D-optimality in experimental design in the sense that D-optimal designs minimize the area of the F distribution confidence ellipsoid for $\boldsymbol{b}$ (see Atkinson, Donev and Tobias, 2007). This relation is outlined in Section 2.4.3. Confidence sets C_2, C_3 and C_h for the two-sided two-segment, three-segment and hyperbolic bands respectively can be generated from the respective confidence regions R_2, R_3 and R_h via the same linear transformation $U^{-1}(\hat{\boldsymbol{b}}-\boldsymbol{b})/\hat{\sigma}$. Subsequently, comparisons among the three types of confidence bands reduce to comparisons among the areas of R_2, R_3 and R_h. Using MACS criterion for simple linear regression, Liu and Hayter (2007) have shown that if the whole range of covariate is of interest, then the hyperbolic band is the recommended band among the three types of confidence bands. Furthermore, if the range of interest is finite, then a restricted hyperbolic band might be recommended, although the three-segment band can be preferable in certain cases.

2.4.3 Relation of MACS to D-optimality

D-optimality is the most intensively studied of all design criteria. It is based on the determinant of the information matrix for the design, $|X^T X|$, which is equal to the reciprocal of the determinant of the variance-covariance matrix for the least squares estimates of the linear parameters of the model, $\frac{1}{|X^T X|^{-1}}$. Designs which maximize $|X^T X|$ are called D-optimum designs. They minimize the content of the confidence region for the parameters $\boldsymbol{b}$ of the model. This is shown below for the model (2.1).

Since, $\hat{\boldsymbol{b}} \sim N(\boldsymbol{b}, \sigma^2 (X^T X)^{-1})$, we have $\frac{1}{\sigma}(X^T X)^{\frac{1}{2}}(\hat{\boldsymbol{b}} - \boldsymbol{b}) \sim N(0, I_2)$ so that

$$\left\{\frac{1}{\sigma}(X^T X)^{\frac{1}{2}}(\hat{\boldsymbol{b}} - \boldsymbol{b})\right\}^T \left\{\frac{1}{\sigma}(X^T X)^{\frac{1}{2}}(\hat{\boldsymbol{b}} - \boldsymbol{b})\right\} \sim \chi^2_2$$

$$\frac{1}{\sigma^2}(\hat{\boldsymbol{b}} - \boldsymbol{b})^T (X^T X)(\hat{\boldsymbol{b}} - \boldsymbol{b}) \sim \chi^2_2$$

$$\frac{(\hat{\boldsymbol{b}} - \boldsymbol{b})^T (X^T X)(\hat{\boldsymbol{b}} - \boldsymbol{b})}{2\sigma^2} \sim \frac{\chi^2_2}{2}$$

$$\frac{(\hat{\boldsymbol{b}} - \boldsymbol{b})^T (X^T X)(\hat{\boldsymbol{b}} - \boldsymbol{b})}{2\hat{\sigma}^2} \sim \frac{\left(\frac{\chi^2_2}{2}\right)}{\left(\frac{\chi^2_\nu}{\nu}\right)} = F_{2,\nu}.$$

So, $P\left\{(\hat{\boldsymbol{b}} - \boldsymbol{b})^T (X^T X)(\hat{\boldsymbol{b}} - \boldsymbol{b}) \leq 2\hat{\sigma}^2 F^{\alpha}_{2,\nu}\right\} = 1-\alpha$ and the region

$$R_{\boldsymbol{b}} = \left\{\boldsymbol{b} : (\hat{\boldsymbol{b}} - \boldsymbol{b})^T (X^T X)(\hat{\boldsymbol{b}} - \boldsymbol{b}) \leq 2\hat{\sigma}^2 F_{2,\nu}\right\}$$

is therefore a confidence region for $\boldsymbol{b}$ and it takes the form of an ellipse (ellipsoid for multiple linear regression). The area of $R_{\boldsymbol{b}}$ is given by

$$\begin{aligned} Area(R_{\boldsymbol{b}}) &= \iint_{R_{\boldsymbol{b}}} 1\, d\boldsymbol{b} \\ &= \iint_{(\hat{\boldsymbol{b}}-\boldsymbol{b})^T(X^TX)(\hat{\boldsymbol{b}}-\boldsymbol{b}) \,\leq\, 2\hat{\sigma}^2F_{2,\nu}} 1\, d\boldsymbol{b} \\ &= \iint_{\boldsymbol{\omega}^T\boldsymbol{\omega} \,\leq\, 2\hat{\sigma}^2F_{2,\nu}} |(X^TX)^{-\frac{1}{2}}|\, d\boldsymbol{\omega} \qquad \Big(\text{where } \boldsymbol{\omega} = (X^TX)^{\frac{1}{2}}(\hat{\boldsymbol{b}}-\boldsymbol{b})\Big) \\ &= \frac{1}{\sqrt{|X^TX|}} \iint_{\boldsymbol{\omega}^T\boldsymbol{\omega} \,\leq\, 2\hat{\sigma}^2F_{2,\nu}} 1\, d\boldsymbol{\omega} \\ &= \frac{2\pi\hat{\sigma}^2F_{2,\nu}}{\sqrt{|X^TX|}} \end{aligned}$$

since $\boldsymbol{\omega}^T\boldsymbol{\omega} = 2\sigma^2F_{2,\nu}$ defines a circle with radius $\sqrt{2\sigma^2F_{2,\nu}}$. Therefore, maximizing $\sqrt{|X^TX|}$ will minimize the area of the confidence region for $\boldsymbol{b}$. Optimality under MACS criterion is also achieved by finding the smallest area of confidence set for $\boldsymbol{b}$.

2.4.4 Optimal design for simultaneous confidence bands under MACS criterion

Liu and Hayter (2007) carried out comparisons among confidence bands in simple linear regression when the design and the range were given. With a chosen $1-\alpha$ confidence level, they concluded that the best band over the whole line is Scheffé's band. When the range of interest is finite, they found that whether the hyperbolic band or the three-segment band is better depends on the value of ϕ. When ϕ is large, the hyperbolic band is better, whereas when ϕ is small, the three-segment band is preferable. Moreover, they also showed that the best two-segment and three-segment bands are given by $c_{2,2,1} = c_{2,2,2}$ and $c_{3,2,1} = c_{3,2,2}$ respectively, that is when they have equal critical constants.

Besides using MACS criterion to compare confidence bands, the criterion can also be used to find the experimental design that leads to the minimum area of the corresponding confidence sets. This was shown by Atkinson, Donev and Tobias (2007) in their example for Scheffé's band which is based on a regression model that holds for the whole real line. However, in most problems, a regression model holds only over a finite interval of the covariate and thus, a confidence band over a finite interval is of interest. The optimization problem for a simultaneous confidence band over a finite range of interest is therefore to find the design that minimizes the area of the corresponding confidence set.

Two-sided two-segment bands

The confidence set corresponding to the two-sided two-segment band, C_2, is given by

$$C_2 = \left\{ \boldsymbol{b} :\; U^{-1}\frac{(\hat{\boldsymbol{b}}-\boldsymbol{b})}{\hat{\sigma}} \in R_2 \right\}$$

which satisfies

$$P\{\boldsymbol{b} \in C_2\} = P\{\boldsymbol{T} \in R_2\} = 1-\alpha$$

and can also be expressed as a linear transformation of R_2:

$$C_2 = \{\boldsymbol{b} : \ \boldsymbol{b} \in \hat{\boldsymbol{b}} + \hat{\sigma} U R_2\}.$$

From Section 2.3.1, R_2 was found to be given by a rectangular region. The area of that region can be readily calculated to be $4c_{2,2,1}c_{2,2,2}$. Hence, the area of C_2 is given by

$$\begin{aligned}
Area(C_2) &= \iint_{C_2} 1 \, d\boldsymbol{b} \\
&= \iint_{\boldsymbol{b} \, \in \, \hat{\boldsymbol{b}} \, + \, U(\hat{\sigma} R_2)} 1 \, d\boldsymbol{b} \\
&= \iint_{\boldsymbol{\omega} \, \in \, \hat{\sigma} R_2} |U| \, d\boldsymbol{\omega} \qquad \left(\text{where } \boldsymbol{\omega} = U^{-1}(\hat{\boldsymbol{b}} - \boldsymbol{b})\right) \\
&= \iint_{\boldsymbol{\rho} \, \in \, R_2} \hat{\sigma}^2 |U| \, d\boldsymbol{\rho} \qquad \left(\text{where } \boldsymbol{\rho} = \frac{1}{\hat{\sigma}}\boldsymbol{\omega}\right) \\
&= \hat{\sigma}^2 |U| \iint_{\boldsymbol{\rho} \, \in \, R_2} 1 \, d\boldsymbol{\rho} \quad \left(\text{where } \iint_{\boldsymbol{\rho} \, \in \, R_2} 1 \, d\boldsymbol{\rho} \text{ is the area of } R_2\right) \\
&= 4\hat{\sigma}^2 c_{2,2,2} c_{2,2,1} |U|.
\end{aligned}$$

Furthermore, Liu and Hayter (2007) have shown that among all two-sided two-segment bands of the form (2.4) satisfying the confidence level requirement (2.2), the best one under MACS criterion is given uniquely by $c_{2,2,1} = c_{2,2,2}$ and R_2 is a square. In that case,

$$Area(C_2) = 4\hat{\sigma}^2 c_{2,2}^2 |U| \tag{2.12}$$

where $c_{2,2} = c_{2,2,1} = c_{2,2,2}$.

The critical constant $c_{2,2}$ does not depend on the design X, as shown by expression (2.5) for the confidence level of a two-sided two-segment band. Consequently, $Area(C_2)$ monotonically increases as $|U|$ increases, *i.e.* $Area(C_2)$ monotonically decreases as $\sqrt{|X^T X|}$ increases. The best two-sided two-segment band, under MACS criterion, is given by the smallest area of corresponding confidence set and is therefore obtained by maximizing $\sqrt{|X^T X|}$. It can be concluded that, under MACS criterion, a D-optimal design leads to the best two-sided two-segment band.

Two-sided three-segment bands

The confidence set corresponding to the two-sided three-segment band, C_3, is given by

$$C_3 = \left\{ \boldsymbol{b} : \ U^{-1} \frac{(\hat{\boldsymbol{b}} - \boldsymbol{b})}{\hat{\sigma}} \in R_3 \right\}$$

which satisfies

$$P\{\boldsymbol{b} \in C_3\} = P\{\boldsymbol{T} \in R_3\} = 1 - \alpha.$$

From Section 2.3.2, R_3 was found to be given by a region given by a parallelogram. The area of R_3 can be easily calculated to be $4c_{3,2,2}c_{3,2,1}/\sin\phi$. Using a similar linear transformation as above, the area of C_3 is given by

$$Area(C_3) = \frac{4\hat{\sigma}^2 c_{3,2,2} c_{3,2,1} |U|}{\sin\phi}.$$

In this case, the angle ϕ depends on the range of interest (a, A) and the design X. Furthermore, the critical constants depend on the angle ϕ and thus on the design X as well through expression (2.9). Analytical minimization of $Area(C_3)$ is more complicated than that of $Area(C_2)$ and thus, numerical methods are used instead. Liu and Hayter (2007) have shown that among all two-sided three-segment bands of the form (2.7), satisfying the confidence level requirement (2.2), the best one under MACS criterion is given uniquely by $c_{3,2,1} = c_{3,2,2}$ and R_3 is a rhombus. In this case,

$$Area(C_3) = \frac{4\hat{\sigma}^2 c_{3,2}^2 |U|}{\sin\phi} \tag{2.13}$$

where $c_{3,2} = c_{3,2,1} = c_{3,2,2}$. In Section 2.5.4, some numerical exploration is carried out to show that $Area(C_3)$ monotonically decreases as $\sqrt{|X^T X|}$ increases, so that it can be deduced that, under MACS criterion, a D-optimal design leads to the best two-sided three-segment band.

Two-sided hyperbolic bands

The confidence set corresponding to the two-sided hyperbolic band, C_h, is given by

$$C_h = \left\{ \boldsymbol{b} : \; U^{-1} \frac{(\hat{\boldsymbol{b}} - \boldsymbol{b})}{\hat{\sigma}} \in R_h \right\}$$

which satisfies

$$P\{\boldsymbol{b} \in C_h\} = P\{\boldsymbol{T} \in R_h\} = 1 - \alpha.$$

From section 2.3.3, R_h was found to be given by a spindle region. The area of the region within the spindle can be calculated to be $c_{h,2}^2 \left[\phi + 2\cot\left(\frac{\phi}{2}\right)\right]$. Consequently,

$$Area(C_h) = \hat{\sigma}^2 c_{h,2}^2 |U| \left[\phi + 2\cot\left(\frac{\phi}{2}\right)\right]. \tag{2.14}$$

In this case as well, the angle ϕ depends on the range of interest (a, A) and the design X. Then, the critical constants depend on the angle ϕ and thus on the design X through expression (2.11). Minimization of $Area(C_h)$ is explored using numerical methods in Section 2.5.4 to show that $Area(C_h)$ monotonically decreases as $\sqrt{|X^T X|}$ increases so that it can be deduced that, under MACS criterion, a D-optimal design also leads to the best two-sided hyperbolic band.

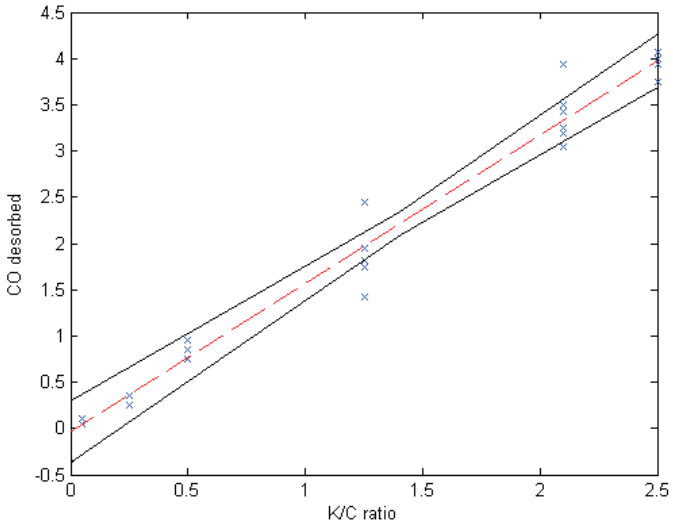
2.5 Numerical examples and exploration

Atkinson, Donev and Tobias (2007) used the Desorption of Carbon Monoxide dataset as their first example of simple linear regression analysis. In the experiment, graphitized carbon was impregnated with potassium carbonate and then heated in a stream of 15% carbon dioxide in nitrogen. The yield was the total amount of carbon monoxide desorbed and it was measured against the initial potassium/carbon ratio. The results of 22 observations are shown in Table 2.1.

Table 2.1: The Desorption of Carbon Monoxide, *Atkinson, Donev and Tobias (2007)*

Observation	Initial K/C atomic ratio (%)	CO absorbed (mole/mole C) (%)
1	0.05	0.05
2	0.05	0.1
3	0.25	0.25
4	0.25	0.35
5	0.5	0.75
6	0.5	0.85
7	0.5	0.95
8	1.25	1.42
9	1.25	1.75
10	1.25	1.82
11	1.25	1.95
12	1.25	2.45
13	2.1	3.05
14	2.1	3.19
15	2.1	3.25
16	2.1	3.43
17	2.1	3.5
18	2.1	3.93
19	2.5	3.75
20	2.5	3.93
21	2.5	3.99
22	2.5	4.07

Figure 2.10: A 0.95 level two-sided two-segment band for desorption CO dataset



The mean of the amount of carbon monoxide desorbed, $\bar{x}$, is equal to 1.4068 mole/mole C %. Modelling the dataset in the form of equation (2.2), $\hat{\boldsymbol{b}} = (\hat{b}_0, \hat{b}_1)^T = (-0.0380, 1.6031)$ and $\hat{\sigma} = 0.0612$

$$(X^T X)^{-1} = \begin{pmatrix} 0.1646 & -0.0847 \\ -0.0847 & 0.0602 \end{pmatrix} \quad \text{and} \quad U = \begin{pmatrix} 0.3779 & -0.1476 \\ -0.1476 & 0.1960 \end{pmatrix}.$$

The confidence level is fixed at $1 - \alpha = 0.95$ and the range of interest is set as $(a, A) = (0, 2.5)$, so that $v(1, a) = 0.1646$, $v(1, A) = 0.1174$, $v(1, \bar{x}) = 0.0455$, $v(0, 1) = 0.0602$ and $\phi = 1.9167$ *rad*. Hence, the simultaneous two-sided two-segment, three-segment and hyperbolic band and their corresponding confidence sets can be constructed for the dataset.

2.5.1 Two-sided two-segment band for the desorption of carbon monoxide dataset

From expression (2.6) and confidence level of $1 - \alpha = 0.95$, $c_{2,2}$ can be evaluated for the dataset to be 2.4109. Hence, expression (2.4) is used to construct the 0.95 level two-sided two-segment band for the desorption of carbon monoxide dataset, as depicted in Figure 2.10, together with the least squares regression line and the 22 observations. The corresponding region R_2 and the confidence set C_2 for $\boldsymbol{b}$ are illustrated in Figure 2.11 and Figure 2.12 respectively. Hence, from equation (2.12), the area of the confidence set, C_2, corresponding to the two-sided two-segment band for the desorption of carbon monoxide dataset is calculated to be 0.0744 $units^2$.

2.5.2 Two-sided three-segment band for the desorption of carbon monoxide dataset

From expression (2.9) and confidence level of $1 - \alpha = 0.95$, $c_{3,2}$ can be evaluated for the dataset to be 2.3970. Hence, expression (2.7) is used to construct the 0.95 level two-

Figure 2.11: The region R_2 for desorption CO dataset

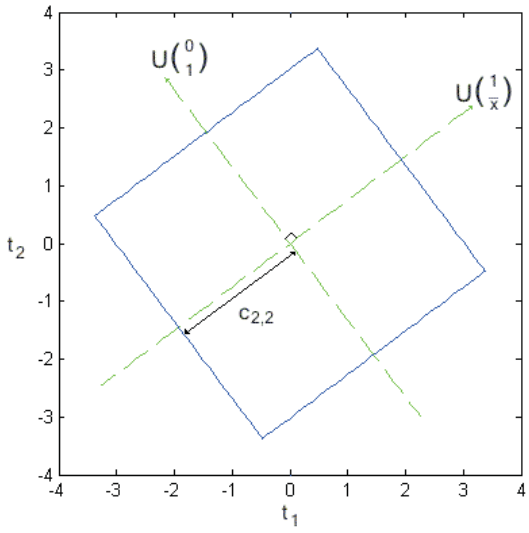


Figure 2.12: The confidence set C_2 for desorption CO dataset

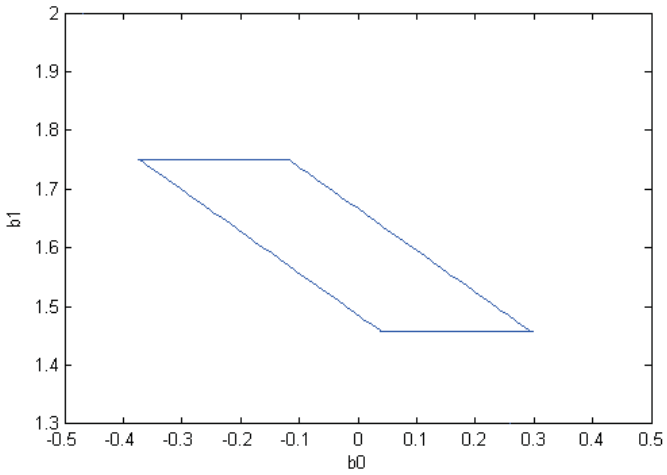
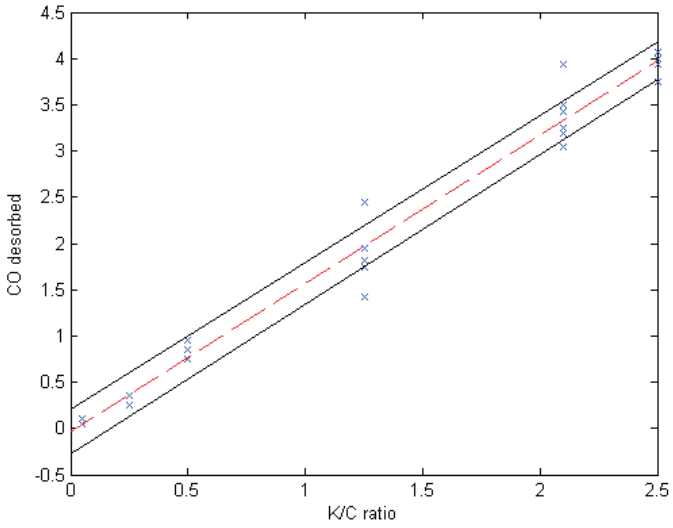


Figure 2.13: A 0.95 level two-sided three-segment band for desorption CO dataset



sided three-segment band for the desorption of carbon monoxide dataset, as shown in Figure 2.13. The corresponding region R_3 and the confidence set C_3 for $\boldsymbol{b}$ are illustrated in Figure 2.14 and Figure 2.15 respectively. Hence, from equation (2.13), the area of the confidence set C_3 for the desorption of carbon monoxide dataset is calculated to be 0.0782 $units^2$. Since this value is larger than that of the two-sided two-segment band, the two-sided two-segment band seems preferable to the two-sided three-segment band for the dataset under the MACS criterion. However, note that the two-segment band is intrinsically defined on the whole range $x \in (-\infty, \infty)$ and cannot be directly compared to the three-segment and hyperbolic bands which are defined on $x \in (a, A) = (0, 2.5)$.

2.5.3 Two-sided hyperbolic band for the desorption of carbon monoxide dataset

From expression (2.11) and confidence level of $1 - \alpha = 0.95$, $c_{h,2}$ can be evaluated for the dataset to be 2.5875. Hence, expression (2.10) is used to construct the 0.95 level two-sided hyperbolic band for the desorption of carbon monoxide dataset, as shown in Figure 2.16. The corresponding region R_h and the confidence set C_h for $\boldsymbol{b}$ are illustrated in Figure 2.17 and Figure 2.18 respectively. Hence, from equation (2.14), the area of the confidence set, C_h, corresponding to the two-sided hyperbolic band for the desorption of carbon monoxide dataset is calculated to be 0.0712 $units^2$. Therefore, under MACS criterion, the 0.95 level two-sided hyperbolic band is the best band for the desorption of carbon monoxide dataset among the three types of 0.95 level simultaneous two-sided confidence bands.

2.5.4 Numerical exploration

In section 2.4.4, an analytical derivation that D-optimal designs lead to the best two-sided three-segment and hyperbolic bands under MACS criterion was not available so far.

Figure 2.14: The region R_3 for desorption CO dataset

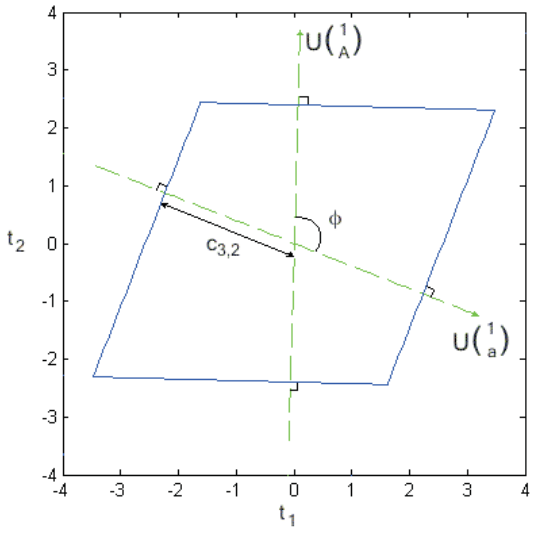


Figure 2.15: The confidence set C_3 for desorption CO dataset

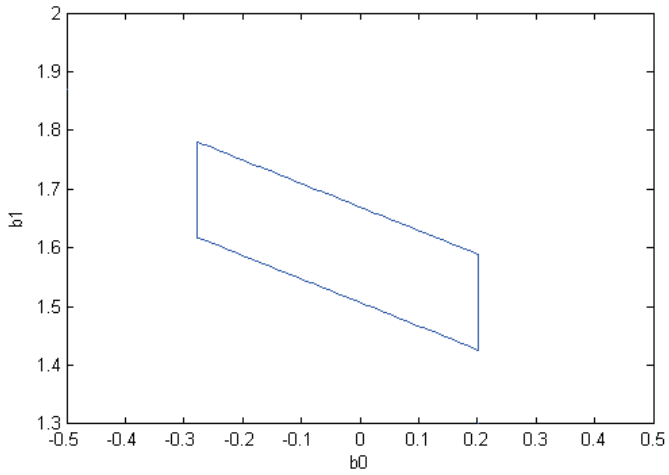


Figure 2.16: A 0.95 level two-sided hyperbolic band for desorption CO dataset

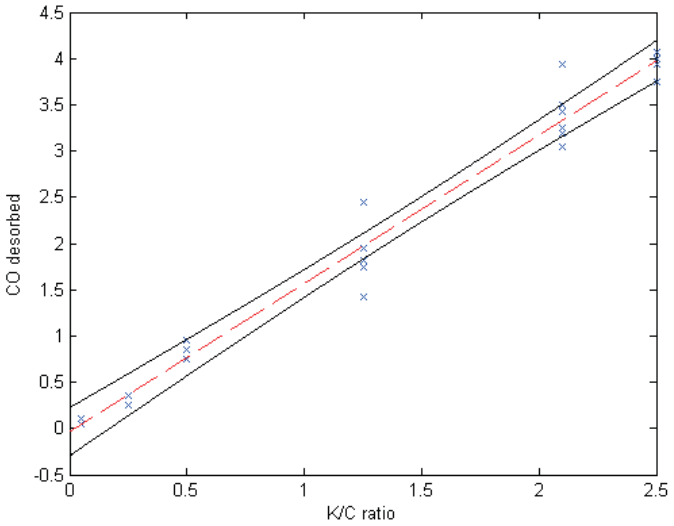


Figure 2.17: The region R_h for desorption CO dataset

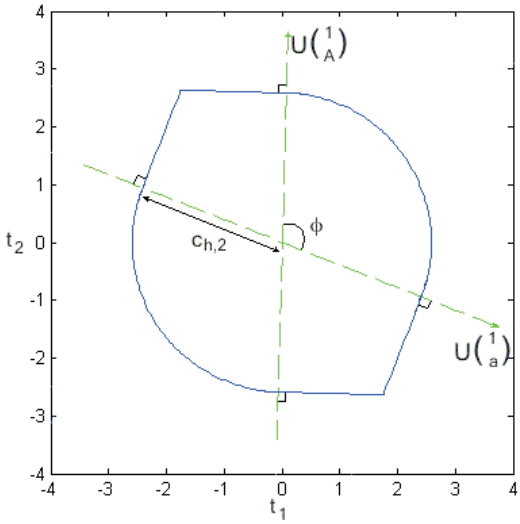
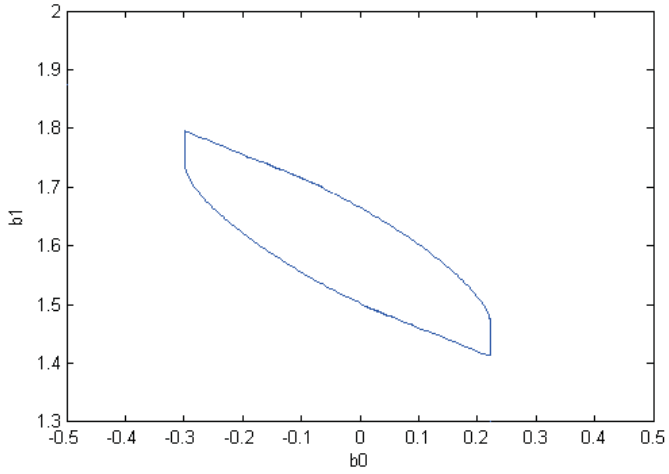


Figure 2.18: The confidence set C_h for desorption CO dataset



Subsequently, numerical methods have been used for the Desorption of Carbon Monoxide dataset, with a range of interest of $(a, A) = (0, 2.5)$ and a confidence level of $1 - \alpha = 0.95$. First of all, a translation in the model (2.1) is necessary to simplify the calculation of U and ϕ independently. The new model is given as

$$y_j = \beta_0 + \beta_1(x_j - \bar{x}) + e_j$$

where $\beta_0 = b_0 + b_1\bar{x}$ and $\beta_1 = b_1$. The range of interest after translation is $(a - \bar{x}, A - \bar{x})$. Hence,

$$(X^TX)^{-1} = \begin{pmatrix} \frac{1}{n} & 0 \\ 0 & \frac{1}{\sum_{i=1}^{n}(x_i-\bar{x})^2} \end{pmatrix} = \begin{pmatrix} \frac{1}{n} & 0 \\ 0 & \frac{1}{s_{xx}} \end{pmatrix}$$

so that

$$U = (X^TX)^{-\frac{1}{2}} = \begin{pmatrix} \frac{1}{\sqrt{n}} & 0 \\ 0 & \frac{1}{\sqrt{s_{xx}}} \end{pmatrix}$$

and

$$|U| = \frac{1}{\sqrt{|X^TX|}} = \frac{1}{\sqrt{ns_{xx}}}.$$

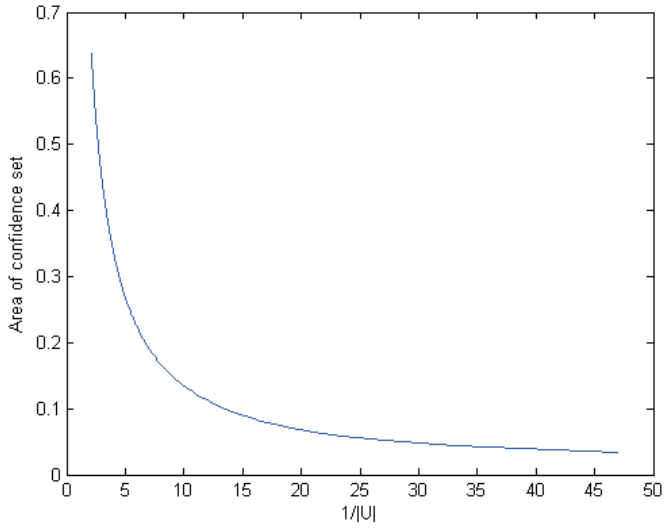
Hence, $U \begin{pmatrix} 1 \\ a - \bar{x} \end{pmatrix} = \left(\frac{1}{\sqrt{n}}, \frac{(a-\bar{x})}{\sqrt{s_{xx}}}\right)^T$ and $U \begin{pmatrix} 1 \\ A - \bar{x} \end{pmatrix} = \left(\frac{1}{\sqrt{n}}, \frac{(A-\bar{x})}{\sqrt{s_{xx}}}\right)^T$.

Moreover,

$$\phi = \cos^{-1}\left[\frac{\left(\frac{1}{n} + \frac{(a-\bar{x})(A-\bar{x})}{s_{xx}}\right)}{\sqrt{\left(\frac{1}{n} + \frac{(a-\bar{x})^2}{s_{xx}}\right)\left(\frac{1}{n} + \frac{(A-\bar{x})^2}{s_{xx}}\right)}}\right]. \tag{2.15}$$

For Desorption of Carbon Monoxide dataset, the value of $\bar{x}$ is fixed. Thus, the input for numerical computation is $s_{xx} = \sum_{i=1}^{n}(x_i - \bar{x})^2$. From s_{xx}, the values of ϕ can be calculated using expression (2.15). The critical constants for the two-sided three-segment band and hyperbolic band can be calculated using expressions (2.9) and (2.11) respectively.

Figure 2.19: Plot of confidence set area against $1/|U| = \sqrt{|X^T X|}$ for best 3-segment band



Hence, the areas of the corresponding confidence sets for the two-sided three-segment and hyperbolic bands can be calculated from expressions (2.13) and (2.14) respectively. Plots of these areas against $1/|U| = \sqrt{|X^T X|}$ are shown in Figure 2.19 and Figure 2.20 for the two-sided three-segment band and hyperbolic band respectively. As $\sqrt{|X^T X|}$ increases, the areas of the corresponding confidence sets are monotonically decreasing in both cases. This supports the deduction made in Section 2.4.4 that a D-optimal design leads to the the best simultaneous two-sided confidence band under MACS criterion for the desorption of carbon monoxide dataset.

In the case where the value of $\bar{x}$ is not fixed, a numerical search for the area of confidence set when both $\bar{x}$ and $1/|U| = \sqrt{|X^T X|}$ vary can be carried out. For instance, a surface of the values of the area of the confidence set corresponding to the best three-segment band when $\bar{x}$ varies within $(a, A) = (0, 2.5)$ and the value of $1/|U| = \sqrt{|X^T X|}$ varies within $(0, 15)$ is depicted in Figure 2.21. It can be observed that the area of confidence set is monotonically decreasing along the direction of the $1/|U|$-axis for a given value of $\bar{x}$.

A similar numerical search is carried out for the hyperbolic band, leading to the plot in Figure 2.22 where the area of confidence set is also monotonically decreasing along the direction of the $1/|U|$-axis for a given value of $\bar{x}$. However, this numerical result can be misleading as $\bar{x}$ and $1/|U|$ cannot be assumed to be restricted within intervals. Thus, although it can be proved analytically that D-optimal designs lead to the best two-segment bands under the MACS criterion, further research is required to show analytically that D-optimal designs lead to the best three-segment and hyperbolic bands under the MACS criterion.

In the next chapter, two new families of simultaneous confidence bands are constructed and compared in simple linear regression, within which optimal confidence bands are identified numerically.

Figure 2.20: Plot of confidence set area against $1/|U| = \sqrt{|X^T X|}$ for hyperbolic band

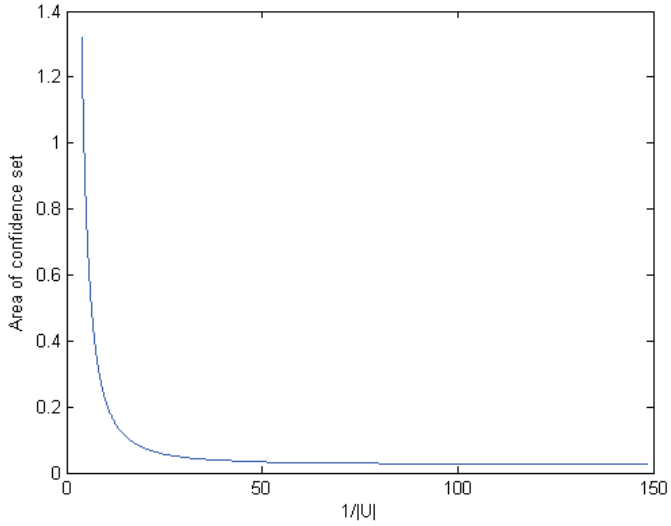


Figure 2.21: Plot of confidence set area against $\bar{x}$ and $1/|U| = \sqrt{|X^T X|}$ for best 3-segment band

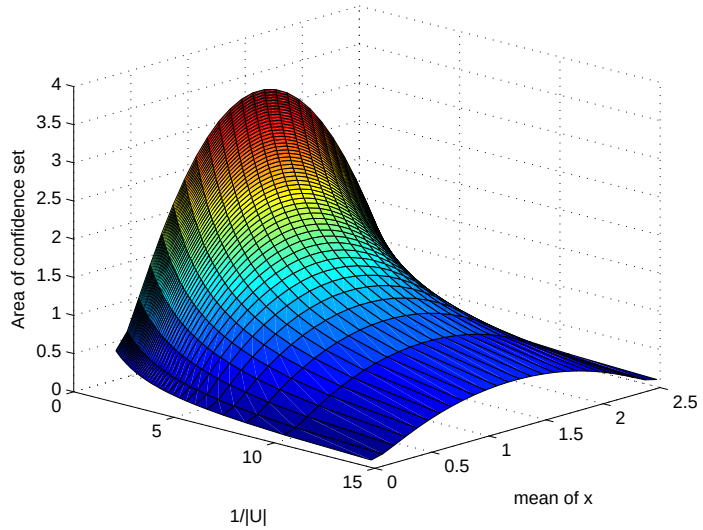
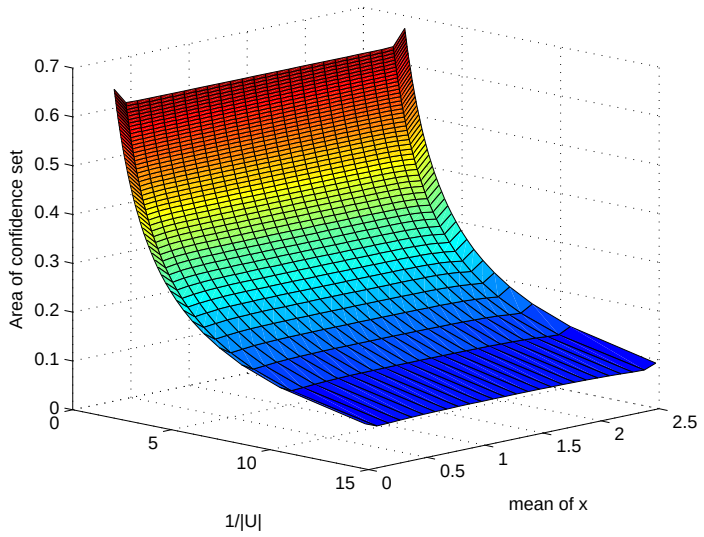


Figure 2.22: Plot of confidence set area against $\bar{x}$ and $1/|U| = \sqrt{|X^T X|}$ for hyperbolic band



Chapter 3

Searching for the best simultaneous confidence band in a particular family of confidence bands in simple linear regression

In this chapter, two families of $1-\alpha$ level confidence bands, which include the hyperbolic band and the best three-segment band as special cases, are defined for $b_0 + b_1 x$. In each family, we search the optimal confidence band under the MACS criterion. The definition of each family is based on a family of confidence sets for $\boldsymbol{b}$ which is in turn defined in terms of a family of sets for $\boldsymbol{T}$ via the transformation $\boldsymbol{T} = U^{-1}(\hat{\boldsymbol{b}} - \boldsymbol{b})/\hat{\sigma}$.

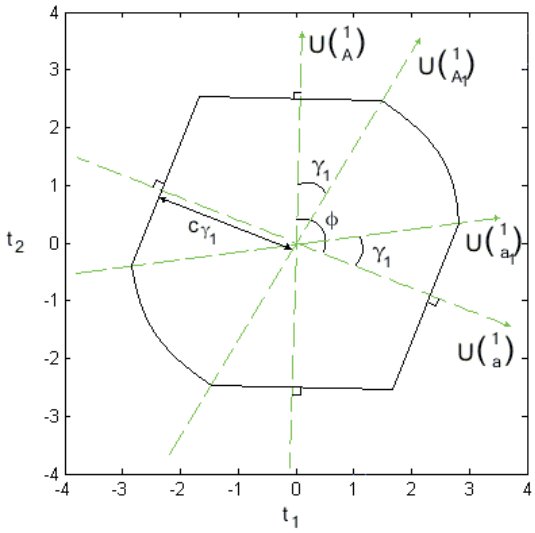
3.1 Family of inner-hyperbolic bands

This family of confidence bands is defined in terms of a family of sets R_{γ_1} for $\boldsymbol{T}$. A set R_{γ_1} for $\boldsymbol{T}$ is defined for each given angle $\gamma_1 \in [0, \phi/2]$, as depicted in Figure 3.1, in the following way. For the given γ_1, the directions $U(1, a_1)^T$ and $U(1, A_1)^T$ marked in Figure 3.1 can be determined uniquely so that $a < a_1 < A_1 < A$ and the angle between $U(1, a)^T$ and $U(1, a_1)^T$ and the angle between $U(1, A_1)^T$ and $U(1, A)^T$ are equal to γ_1. Specifically, $a < a_1 < A_1 < A$ are solved uniquely from

$$\cos\gamma_1 = \frac{\left\{U\begin{pmatrix}1\\a\end{pmatrix}\right\}^T\left\{U\begin{pmatrix}1\\a_1\end{pmatrix}\right\}}{\left\|U\begin{pmatrix}1\\a\end{pmatrix}\right\|\left\|U\begin{pmatrix}1\\a_1\end{pmatrix}\right\|} = \frac{\left\{U\begin{pmatrix}1\\A\end{pmatrix}\right\}^T\left\{U\begin{pmatrix}1\\A_1\end{pmatrix}\right\}}{\left\|U\begin{pmatrix}1\\A\end{pmatrix}\right\|\left\|U\begin{pmatrix}1\\A_1\end{pmatrix}\right\|}. \tag{3.1}$$

The set R_{γ_1} is bounded by a segment of a circle of radius $c_{\gamma_1}/\cos\gamma_1$ centered at the origin between $U(1, a_1)^T$ and $U(1, A_1)^T$ and between $-U(1, a_1)^T$ and $-U(1, A_1)^T$. The remaining boundary of R_{γ_1} is formed by the four line segments which are perpendicular to $U(1, a)^T$ and $U(1, A)^T$ and c_{γ_1} distance in both directions from the origin. By the way of construction, R_{γ_1} is uniquely determined by c_{γ_1} for a given γ_1. We choose c_{γ_1} such that

Figure 3.1: The region R_{γ_1}



$P\{\boldsymbol{T} \in R_{\gamma_1}\} = 1-\alpha$. Note that, for $\gamma_1 = 0$, R_{γ_1} is simply R_h depicted in Figure 2.17 and, for $\gamma_1 = \phi/2$, R_{γ_1} is simply R_3 depicted in Figure 2.14.

The confidence set for $\boldsymbol{b}$ that corresponds to R_{γ_1} is given by

$$C_{\gamma_1} = \left\{ \boldsymbol{b} : \; U^{-1}\frac{(\hat{\boldsymbol{b}} - \boldsymbol{b})}{\hat{\sigma}} \in R_{\gamma_1} \right\}.$$

The confidence band for $b_0 + b_1 x$ that correspond to this confidence set C_{γ_1} can be shown to be given by $\hat{b_0} + \hat{b_1} x \pm \hat{\sigma} H_{\gamma_1}(x)$ with

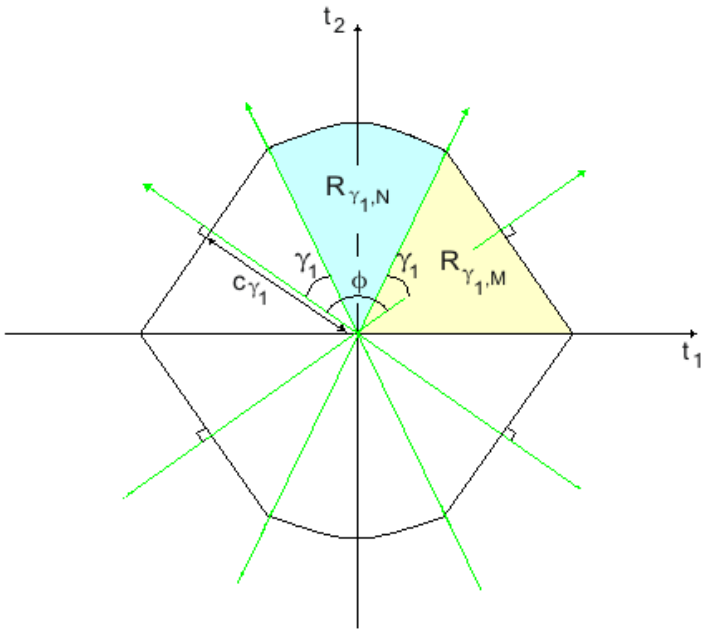
$$H_{\gamma_1}(x) = \begin{cases} \frac{1}{(a_1-a)}\left[(x-a)\frac{c_{\gamma_1}}{\cos\gamma_1}\sqrt{v(1,a_1)} + (a_1-x)c_{\gamma_1}\sqrt{v(1,a)}\right] & \text{for } \forall x \in (a, a_1] \\ \frac{c_{\gamma_1}}{\cos\gamma_1}\sqrt{v(1,x)} & \text{for } \forall x \in (a_1, A_1) \\ \frac{1}{(A-A_1)}\left[(x-A_1)c_{\gamma_1}\sqrt{v(1,A)} + (A-x)\frac{c_{\gamma_1}}{\cos\gamma_1}\sqrt{v(1,A_1)}\right] & \text{for } \forall x \in [A_1, A). \end{cases}$$

It is clear that when γ_1 changes within $[0, \phi/2]$ we have a family of $1-\alpha$ level confidence bands. We call this the family of inner-hyperbolic bands. It is easy to check that, for $\gamma_1 = 0$, this band is just the hyperbolic band and, for $\gamma_1 = \phi/2$, this band is just the best three-segment band.

3.1.1 Confidence level of the inner-hyperbolic band

The critical constant c_{γ_1} of the inner-hyperbolic band is determined from $P\{\boldsymbol{T} \in R_{\gamma_1}\} = 1-\alpha$. Let $R^*_{\gamma_1}$ be the region that is resulted from rotating R_{γ_1} around the origin to the position so that the angle ϕ is divided into two equal halves by the t_2-axis, as depicted in Figure 3.2. Due to the rotational invariance of the probability distribution of $\boldsymbol{T}$, the probability of $\boldsymbol{T}$ in R_{γ_1} is equal to the probability of $\boldsymbol{T}$ in $R^*_{\gamma_1}$. The region $R^*_{\gamma_1}$ is partitioned

Figure 3.2: The region $R^*_{\gamma_1}$



into four triangles of size equal to $R_{\gamma_1,M}$ and two fans of size equal to $R_{\gamma_1,N}$, as illustrated in Figure 3.2. The probability of $\boldsymbol{T}$ in R_{γ_1} is therefore equal to the sum of twice the probability of $\boldsymbol{T}$ in $R_{\gamma_1,N}$ and four times the probability of $\boldsymbol{T}$ in $R_{\gamma_1,M}$.

Furthermore, the region $R_{\gamma_1,M}$ can be expressed as

$$R_{\gamma_1,M} = \left\{ \boldsymbol{T} : \theta \in \left[0, \frac{\pi-\phi}{2} + \gamma_1\right] \ , \ 0 \leq \left(\cos\left(\frac{\pi-\phi}{2}\right), \sin\left(\frac{\pi-\phi}{2}\right)\right) \boldsymbol{T} \leq c_{\gamma_1} \right\}$$

and so

$$\begin{aligned}
& P\{\boldsymbol{T} \in R_{\gamma_1,M}\} \\
= \ & P\left\{ \theta \in \left[0, \frac{\pi-\phi}{2} + \gamma_1\right] \ , \ 0 \leq \left(\cos\left(\frac{\pi-\phi}{2}\right), \sin\left(\frac{\pi-\phi}{2}\right)\right) \boldsymbol{T} \leq c_{\gamma_1} \right\} \\
= \ & \int_0^{\frac{\pi-\phi}{2}+\gamma_1} \frac{1}{2\pi} P\left\{ R \leq \frac{c_{\gamma_1}}{\cos\left(\frac{\pi-\phi}{2} - \theta\right)} \right\} d\theta \\
= \ & \int_0^{\frac{\pi-\phi}{2}+\gamma_1} \frac{1}{2\pi} F_R\left(\frac{c_{\gamma_1}}{\sin(\theta + \frac{\phi}{2})} \right) d\theta \\
= \ & \frac{1}{2\pi}\left[\frac{\pi-\phi}{2} + \gamma_1\right] - \frac{1}{2\pi} \int_0^{\frac{\pi-\phi}{2}+\gamma_1} \left[1 + \frac{c_{\gamma_1}^2}{\nu \sin^2(\theta + \frac{\phi}{2})} \right]^{-\frac{\nu}{2}} d\theta.
\end{aligned}$$

The region $R_{\gamma_1,N}$ can be expressed as

$$R_{\gamma_1,N} = \left\{ \boldsymbol{T} : \theta \in \left[\frac{\pi-\phi}{2} + \gamma_1, \frac{\pi+\phi}{2} - \gamma_1\right] \ , \ \|\boldsymbol{T}\| \leq \frac{c_{\gamma_1}}{\cos\gamma_1} \right\}$$

and so

$$\begin{aligned}
& P\{\boldsymbol{T} \in R_{\gamma_1,N}\} \\
= \; & P\left\{\theta \in \left[\frac{\pi-\phi}{2}+\gamma_1, \frac{\pi+\phi}{2}-\gamma_1\right] \;,\; \|\boldsymbol{T}\| \leq \frac{c_{\gamma_1}}{\cos\gamma_1}\right\} \\
= \; & \frac{\phi-2\gamma_1}{2\pi} P\{R \leq \frac{c_{\gamma_1}}{\cos\gamma_1}\} \\
= \; & \frac{\phi-2\gamma_1}{2\pi} F_R(\frac{c_{\gamma_1}}{\cos\gamma_1}) \\
= \; & \frac{\phi-2\gamma_1}{2\pi}\left[1-\left(1+\frac{c_{\gamma_1}^2}{\nu\cos\gamma_1}\right)^{-\frac{\nu}{2}}\right].
\end{aligned}$$

Therefore, we have

$$\begin{aligned}
& P\{\boldsymbol{T} \in R_{\gamma_1}\} = 4P\{\boldsymbol{T} \in R_{\gamma_1,M}\} + 2P\{\boldsymbol{T} \in R_{\gamma_1,N}\} \\
= \; & \frac{\phi-2\gamma_1}{\pi}\left[1-\left(1+\frac{c_{\gamma_1}^2}{\nu\cos\gamma_1}\right)^{-\frac{\nu}{2}}\right] + \frac{2}{\pi}\left[\frac{\pi-\phi}{2}+\gamma_1\right] \\
- \; & \frac{2}{\pi}\int_0^{\frac{\pi-\phi}{2}+\gamma_1}\left(1+\frac{c_{\gamma_1}^2}{\nu\sin^2(\theta+\frac{\phi}{2})}\right)^{-\frac{\nu}{2}} d\theta. \qquad (3.2)
\end{aligned}$$

Expression (3.2) gives the confidence level of a two-sided inner-hyperbolic band for a given c_{γ_1}, which can be used to calculate the critical constant c_{γ_1} for a given α. When $\gamma_1 = 0$, it matches the confidence level of the two-sided hyperbolic band given by expression (2.11), whereas when $\gamma_1 = \phi/2$, it matches the confidence level of the best two-sided three-segment band given by expression (2.9) when $c_{3,2,1} = c_{3,2,2}$ (see Liu and Hayter, 2007).

3.1.2 Area of confidence set corresponding to the inner-hyperbolic band

From the relationship between C_{γ_1} and R_{γ_1}, similar derivation as in Section 2.4.4 can be used to show that $Area(C_{\gamma_1}) = \hat{\sigma}^2|U|Area(R_{\gamma_1})$. Hence, from Figure 3.1, it is also clear that the region R_{γ_1} can be partitioned into

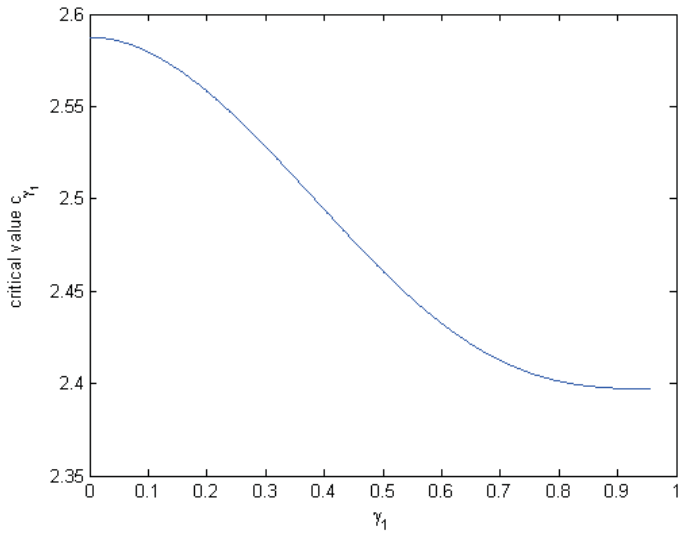
- two fans formed by the lines $U(1,a_1)^T$, $U(1,A_1)^T$ and the boundary of the region R_{γ_1}.
- four small right-angled triangles such as the one formed by $U(1,a)^T$, $U(1,a_1)^T$ and the boundary of the region R_{γ_1}.
- four big right-angled triangles such as the one formed by $U(1,a)^T$, the boundary of the region R_{γ_1} and the line joining the vertices of the region R_{γ_1}.

Therefore, the area of the region R_{γ_1} can be calculated as the sum of twice the area of one of the fans, four times the area of one of the small right-angled triangles and four times the area of one of the big right-angled triangles.

The area of one of the fans is $\frac{1}{2}\frac{c_{\gamma_1}^2}{\cos^2\gamma_1}(\phi-2\gamma_1)$.

The area of one of the small right-angled triangles is $\frac{1}{2}c_{\gamma_1}^2\tan\gamma_1$.

Figure 3.3: Plot of c_{γ_1} against γ_1



The area of one of the big right-angled triangles is $\frac{1}{2}c_{\gamma_1}^2 \cot(\frac{\phi}{2})$.

Therefore, the area of the confidence region R_{γ_1} is given by

$$Area(R_{\gamma_1}) = 2\frac{c_{\gamma_1}^2}{\cos^2\gamma_1}(\phi - 2\gamma_1) + 2c_{\gamma_1}^2 \tan\gamma_1 + 2c_{\gamma_1}^2 \cot(\frac{\phi}{2}).$$

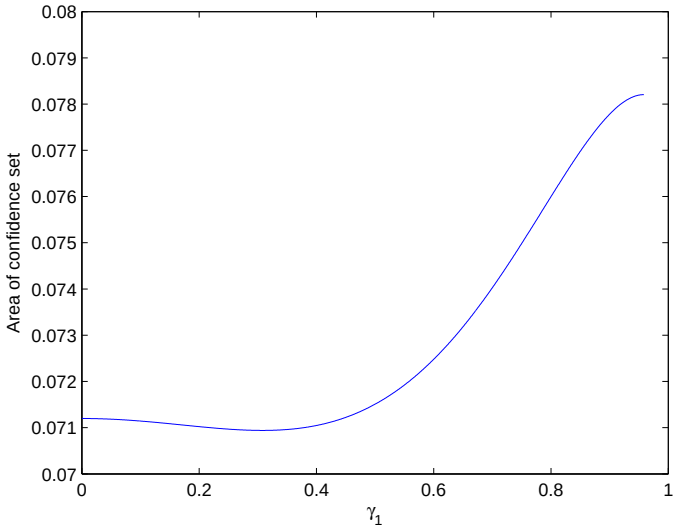
When $\gamma_1 = 0$, $Area(R_{\gamma_1}) = c_{\gamma_1}^2 \left[\phi + 2\cot\left(\frac{\phi}{2}\right)\right]$, which is just the area of the region R_h for the hyperbolic band and when $\gamma_1 = \frac{\phi}{2}$, $Area(R_{\gamma_1}) = 4c_{\gamma_1}^2/\sin\phi$, which is just the area of the region R_3 for the best three-segment band (see Liu and Hayter, 2007). Hence,

$$Area(C_{\gamma_1}) = \hat{\sigma}^2|U| \left[2\frac{c_{\gamma_1}^2}{\cos^2\gamma_1}(\phi - 2\gamma_1) + 2c_{\gamma_1}^2 \tan\gamma_1 + 2c_{\gamma_1}^2 \cot(\frac{\phi}{2})\right]. \qquad (3.3)$$

3.1.3 Searching for the best inner-hyperbolic band

For given ϕ, ν and α, we can numerically search for the best band in the family of $1-\alpha$ level inner-hyperbolic bands by finding the angle $\gamma_1 \in [0, \phi/2]$ that minimizes $Area(C_{\gamma_1})$ under the constraint $P\{\boldsymbol{T} \in R_{\gamma_1}\} = 1-\alpha$. The dataset on the desorption of carbon monoxide from Atkinson, Donev and Tobias (2007) given in Table 3.1 is used to illustrate the numerical search. As in Section 2.5, the regression line $b_0 + b_1 x$ is to be bounded over the range of interest $x \in (a, A) = (0, 2.5)$ by using a $1-\alpha = 0.95$ level simultaneous confidence band. The value of ϕ can be calculated using equation (2.8) to be 1.9167 *rad*. For each value of $\gamma_1 \in [0, \phi/2]$, the corresponding critical constant c_{γ_1} of the inner-hyperbolic band is computed using expression (3.2). A plot of c_{γ_1} against γ_1 is provided in Figure 3.3. Then, the corresponding area of confidence set is calculated from expression (3.3). The area of the confidence set against γ_1 is plotted in Figure 3.4 from which the $\gamma_1 \in [0, \phi/2]$ that gives the MACS, *i.e.* the best inner-hyperbolic band, can be identified. Specifically, the area of the confidence set corresponding to the best inner-hyperbolic band is 0.07094 $units^2$, whereas those corresponding to the hyperbolic and best three-segment

Figure 3.4: Plot of $Area(C_{\gamma_1})$ against γ_1



bands are 0.07120 and 0.07820 $units^2$ respectively. Furthermore, the optimal $\gamma_1 \in [0, \phi/2]$ is given by 0.3076 rad and the critical value $c_{\gamma_1} = 2.5259$. Using equation (3.1) and this optimal value of γ_1, the range (a_1, A_1) is found to be $(0.6601, 1.9906)$. The best inner-hyperbolic band is shown in Figure 3.5 together with the least squares regression line, the 22 observations and the "inner-range" (a_1, A_1) as vertical dashed lines.

3.1.4 Comparisons

The best inner-hyperbolic band can be compared with the best three-segment band and with the hyperbolic band by looking at

$$e_{\phi,3} = \frac{Area(C_I^*)}{Area(C_3)} \quad \text{and} \quad e_{\phi,h} = \frac{Area(C_I^*)}{Area(C_h)}$$

where C_I^* denotes the confidence set of the best inner-hyperbolic band. As a function of $\phi \in (0, \pi)$, $e_{\phi,3}$ is equal to one for $\phi \in (0, \phi^*)$ where ϕ^* is a value depending on ν and α. Then, $e_{\phi,3}$ strictly decreases to zero for $\phi \in (\phi^*, \pi)$. Table 3.1 provides the value of ϕ^* for some combinations of ν and α, while Figure 3.6 presents a typical picture of $e_{\phi,3}$.

From this, it can be concluded that the best inner-hyperbolic band is actually given by the best three-segment band for $\phi \in (0, \phi^*)$, but a more efficient band than the best three-segment band can be found for $\phi \in (\phi^*, \pi)$. The best three-segment band is very in-efficient under the MACS criterion relative to the best inner-hyperbolic band when ϕ is close to π since $e_{\phi,3} \to 0$ as $\phi \to \pi$.

The function $e_{\phi,h}$ first strictly decreases from one and then strictly increases and approaches one over $\phi \in (0, \pi)$. The minimum value of $e_{\phi,h}$ is only marginally smaller than one. A typical picture of $e_{\phi,h}$ is given in Figure 3.7. From this, it can be concluded that the best inner-hyperbolic band is always more efficient than the hyperbolic band under

Figure 3.5: The best 0.95 level inner-hyperbolic band for the desorption CO dataset

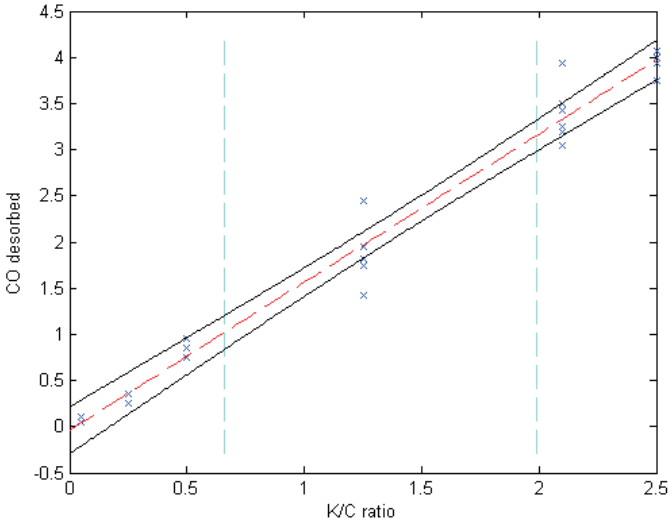


Table 3.1: Values of ϕ^* in *rad* for combinations of $\alpha = 0.01,\ 0.05,\ 0.10$ and $\nu = 10,\ 30,\ \infty$

	$\alpha = 0.10$	$\alpha = 0.05$	$\alpha = 0.01$
$\nu = 10$	0.5240	0.9692	1.0128
$\nu = 30$	0.6261	1.0515	0.99997
$\nu = \infty$	1.0904	1.0483	0.96635

Figure 3.6: Plot of $e_{\phi,3}$ against ϕ

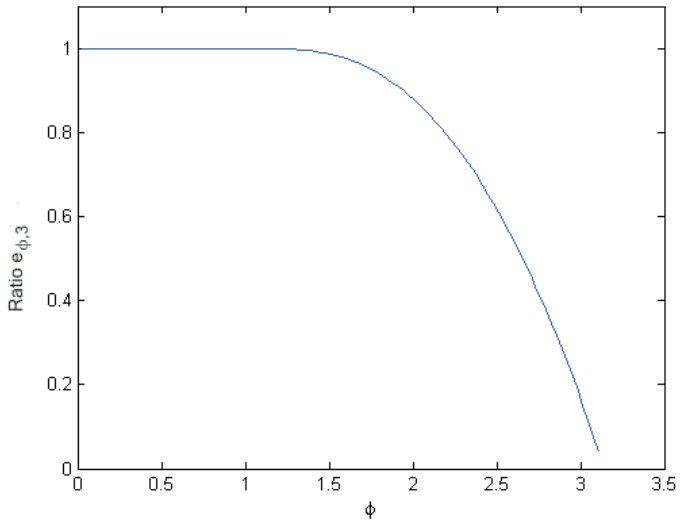


Figure 3.7: Plot of $e_{\phi,h}$ against ϕ

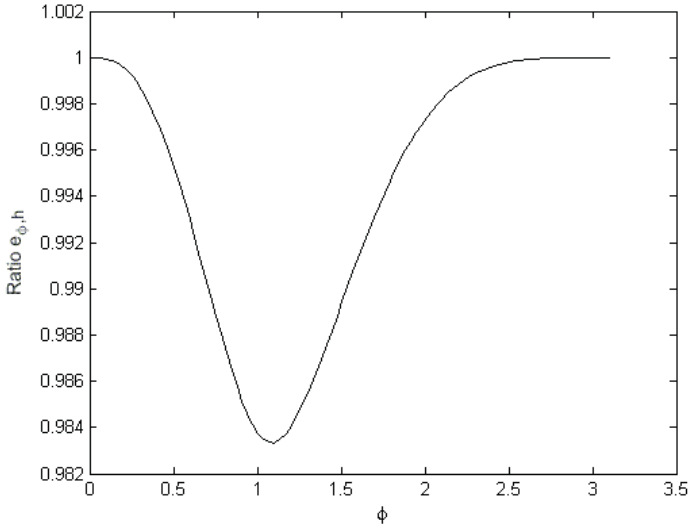
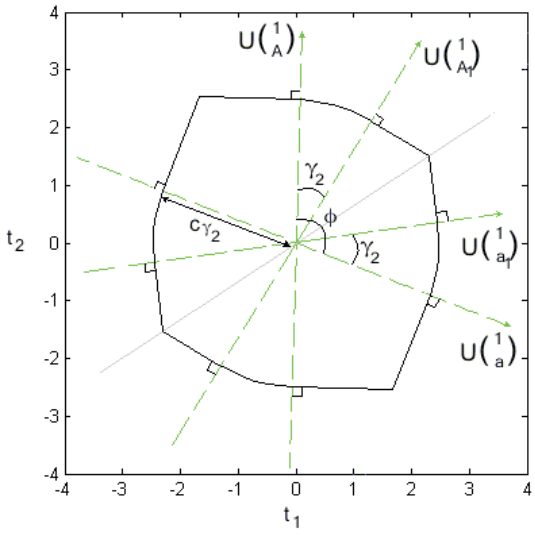


Figure 3.8: The region R_{γ_2}



the MACS criterion, but the gain in efficiency of the best inner-hyperbolic band over the hyperbolic band is never large.

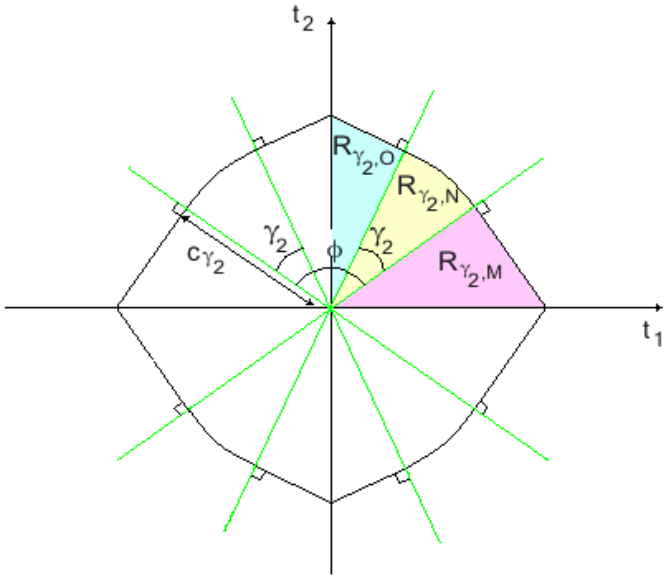
3.2 Family of outer-hyperbolic bands

This family of confidence bands is defined in terms of a family of sets R_{γ_2} for $\boldsymbol{T}$. For each given angle $\gamma_2 \in [0, \phi/2]$, a set R_{γ_2} for $\boldsymbol{T}$ is defined in the following way and depicted in Figure 3.8. For the given γ_2, the directions $U(1, a_1)^T$ and $U(1, A_1)^T$ in Figure 3.8 are determined uniquely so that $a < a_1 < A_1 < A$ and the angles between $U(1, a)^T$ and $U(1, a_1)^T$ and between $U(1, A_1)^T$ and $U(1, A)^T$ are equal to γ_2. The set R_{γ_2} is bounded by a segment of a circle of radius c_{γ_2} centered at the origin between $U(1, a)^T$ and $U(1, a_1)^T$, between $-U(1, a)^T$ and $-U(1, a_1)^T$, between $U(1, A_1)^T$ and $U(1, A)^T$ and between $-U(1, A_1)^T$ and $-U(1, A)^T$. The remaining boundary of R_{γ_2} is formed by the eight line segments that are perpendicular to directions $U(1, a)^T$, $U(1, a_1)^T$, $U(1, A_1)^T$ and $U(1, A)^T$ and c_{γ_2} distance in both directions from the origin. It is clear from this construction that R_{γ_2} is uniquely determined by c_{γ_2} for a given angle γ_2. We choose c_{γ_2} such that $P\{\boldsymbol{T} \in R_{\gamma_2}\} = 1 - \alpha$. Note that, for $\gamma_2 = \phi/2$, R_{γ_2} is simply R_h depicted in Figure 2.17 and, for $\gamma_2 = 0$, R_{γ_2} is simply R_3 depicted in Figure 2.14.

The confidence set for $\boldsymbol{b}$ that corresponds to R_{γ_2} is given by

$$C_{\gamma_2} = \left\{ \boldsymbol{b} : \; U^{-1} \frac{(\hat{\boldsymbol{b}} - \boldsymbol{b})}{\hat{\sigma}} \in R_{\gamma_2} \right\}.$$

Figure 3.9: The region $R^*_{\gamma_2}$



The confidence band for $b_0 + b_1 x$ that correspond to this confidence set C_{γ_2} can be shown to be given by $\hat{b_0} + \hat{b_1} x \pm \hat{\sigma} H_{\gamma_2}(x)$ with

$$H_{\gamma_2}(x) = \begin{cases} c_{\gamma_2}\sqrt{v(1,x)} & \text{for } \forall x \in (a, a_1] \\ \frac{1}{(A_1 - a_1)}\left[(x-a_1)c_{\gamma_2}\sqrt{v(1,A_1)} + (A_1 - x)c_{\gamma_2}\sqrt{v(1,a_1)}\right] & \text{for } \forall x \in (a_1, A_1) \\ c_{\gamma_2}\sqrt{v(1,x)} & \text{for } \forall x \in [A_1, A). \end{cases}$$

The confidence level of the band is $1-\alpha$ since $P\{\boldsymbol{T} \in R_{\gamma_2}\} = 1-\alpha$. It is clear that when γ_2 changes within $[0, \phi/2]$ we have a family of $1-\alpha$ level confidence bands which is called the family of outer-hyperbolic bands. It is easy to check that, for $\gamma_2 = 0$, the band is just the best three-segment band and, for $\gamma_2 = \phi/2$, the band is just the hyperbolic band.

3.2.1 Confidence level of the outer-hyperbolic band

To calculate c_{γ_2} from $P\{\boldsymbol{T} \in R_{\gamma_2}\} = 1-\alpha$, we derive an expression for $P\{\boldsymbol{T} \in R_{\gamma_2}\}$. Let $R^*_{\gamma_2}$ be the region that is resulted from rotating R_{γ_2} around the origin to the position so that the angle ϕ is divided into two equal halves by the t_2-axis, as depicted in Figure 3.9. Due to the rotational invariance of the probability distribution of $\boldsymbol{T}$, the probability of $\boldsymbol{T}$ in R_{γ_2} is equal to the probability of $\boldsymbol{T}$ in $R^*_{\gamma_2}$. The region $R^*_{\gamma_2}$ is partitioned into four triangles of equal size to $R_{\gamma_2,M}$, four fans of size equal to $R_{\gamma_2,N}$ and four triangles of size equal to $R_{\gamma_2,O}$, which are illustrated in Figure 3.9. The probability of $\boldsymbol{T}$ in R_{γ_2} is therefore

equal to the sum of four times the probability of $\boldsymbol{T}$ in $R_{\gamma_2,O}$, four times the probability of $\boldsymbol{T}$ in $R_{\gamma_2,M}$ and four times the probability of $\boldsymbol{T}$ in $R_{\gamma_2,N}$.

Furthermore, the region $R_{\gamma_2,M}$ can be expressed as

$$R_{\gamma_2,M} = \left\{ \boldsymbol{T} : \theta \in \left[0, \frac{\pi-\phi}{2}\right] \ , \ 0 \leq \left(\cos\left(\frac{\pi-\phi}{2}\right), \sin\left(\frac{\pi-\phi}{2}\right)\right) \boldsymbol{T} \leq c_{\gamma_2} \right\}$$

and so

$$\begin{aligned}
& P\{\boldsymbol{T} \in R_{\gamma_2,M}\} \\
= \ & P\left\{ \theta \in \left[0, \frac{\pi-\phi}{2}\right] \ , \ 0 \leq \left(\cos\left(\frac{\pi-\phi}{2}\right), \sin\left(\frac{\pi-\phi}{2}\right)\right) \boldsymbol{T} \leq c_{\gamma_2} \right\} \\
= \ & \int_0^{\frac{\pi-\phi}{2}} \frac{1}{2\pi} P\left\{ R \leq \frac{c_{\gamma_2}}{\cos\left(\frac{\pi-\phi}{2} - \theta\right)} \right\} d\theta \\
= \ & \int_0^{\frac{\pi-\phi}{2}} \frac{1}{2\pi} F_R\left(\frac{c_{\gamma_2}}{\sin(\theta + \frac{\phi}{2})} \right) d\theta \\
= \ & \frac{1}{2\pi}\left[\frac{\pi-\phi}{2}\right] - \frac{1}{2\pi} \int_0^{\frac{\pi-\phi}{2}} \left[1 + \frac{c_{\gamma_2}^2}{\nu \sin^2(\theta + \frac{\phi}{2})} \right]^{-\frac{\nu}{2}} d\theta.
\end{aligned}$$

The region $R_{\gamma_2,N}$ can be expressed as

$$R_{\gamma_2,N} = \left\{ \boldsymbol{T} : \theta \in \left[\frac{\pi-\phi}{2}, \frac{\pi-\phi}{2} + \gamma_2\right] \ , \ \|\boldsymbol{T}\| \leq c_{\gamma_2} \right\}$$

and so

$$\begin{aligned}
& P\{\boldsymbol{T} \in R_{\gamma_2,N}\} \\
= \ & P\left\{ \theta \in \left[\frac{\pi-\phi}{2}, \frac{\pi-\phi}{2} + \gamma_2\right] \ , \ \|\boldsymbol{T}\| \leq c_{\gamma_2} \right\} \\
= \ & \frac{\gamma_2}{2\pi} P\{R \leq c_{\gamma_2}\} \\
= \ & \frac{\gamma_2}{2\pi} F_R(c_{\gamma_2}) \\
= \ & \frac{\gamma_2}{2\pi}\left[1 - \left(1 + \frac{c_{\gamma_2}^2}{\nu} \right)^{-\frac{\nu}{2}} \right].
\end{aligned}$$

The region $R_{\gamma_2,O}$ can be expressed as

$$R_{\gamma_2,O} = \left\{ \boldsymbol{T} : \theta \in \left[\frac{\pi-\phi}{2} + \gamma_2, \frac{\pi}{2}\right] \ , \ 0 \leq \left(\sin\left(\frac{\phi-2\gamma_2}{2}\right), \cos\left(\frac{\phi-2\gamma_2}{2}\right)\right) \boldsymbol{T} \leq c_{\gamma_2} \right\}$$

and so

$$\begin{aligned}
& P\{\boldsymbol{T} \in R_{\gamma_2,O}\} \\
= \ & P\left\{ \theta \in \left[\frac{\pi-\phi}{2} + \gamma_2, \frac{\pi}{2}\right] \ , \ 0 \leq \left(\sin\left(\frac{\phi-2\gamma_2}{2}\right), \cos\left(\frac{\phi-2\gamma_2}{2}\right)\right) \boldsymbol{T} \leq c_{\gamma_2} \right\} \\
= \ & \int_{\frac{\pi-\phi}{2}+\gamma_2}^{\frac{\pi}{2}} \frac{1}{2\pi} P\left\{ R \leq \frac{c_{\gamma_2}}{\cos\left[\theta - \left(\frac{\pi-\phi}{2} + \gamma_2\right)\right]} \right\} d\theta \\
= \ & \int_{\frac{\pi-\phi}{2}+\gamma_2}^{\frac{\pi}{2}} \frac{1}{2\pi} F_R\left(\frac{c_{\gamma_2}}{\cos\left[\theta - \left(\frac{\pi-\phi}{2} + \gamma_2\right)\right]} \right) d\theta \\
= \ & \frac{\phi-2\gamma_2}{4\pi} - \frac{1}{2\pi} \int_{\frac{\pi-\phi}{2}+\gamma_2}^{\frac{\pi}{2}} \left[1 + \frac{c_{\gamma_2}^2}{\nu \cos^2(\left[\theta - \left(\frac{\pi-\phi}{2} + \gamma_2\right)\right])} \right]^{-\frac{\nu}{2}} d\theta.
\end{aligned}$$

Therefore, we have

$$
\begin{aligned}
& P\{\boldsymbol{T} \in R_{\gamma_2}\} \\
= \quad & 4P\{\boldsymbol{T} \in R_{\gamma_2,M}\} + 4P\{\boldsymbol{T} \in R_{\gamma_2,N}\} + 4P\{\boldsymbol{T} \in R_{\gamma_2,O}\} \\
= \quad & \frac{2}{\pi}\left[\frac{\pi-\phi}{2}\right] - \frac{2}{\pi}\int_0^{\frac{\pi-\phi}{2}} \left[1 + \frac{c_{\gamma_2}^2}{\nu \sin^2(\theta+\frac{\phi}{2})}\right]^{-\frac{\nu}{2}} d\theta \\
+ \quad & \frac{2\gamma_2}{\pi}\left[1 - \left(1 + \frac{c_{\gamma_2}^2}{\nu}\right)^{-\frac{\nu}{2}}\right] + \frac{\phi - 2\gamma_2}{\pi} \\
- \quad & \frac{2}{\pi}\int_{\frac{\pi-\phi}{2}+\gamma_2}^{\frac{\pi}{2}} \left[1 + \frac{c_{\gamma_2}^2}{\nu \cos^2(\left[\theta - \left(\frac{\pi-\phi}{2}+\gamma_2\right)\right])}\right]^{-\frac{\nu}{2}} d\theta. \qquad (3.4)
\end{aligned}
$$

Expression (3.4) gives the confidence level of a two-sided outer-hyperbolic band for a given c_{γ_2}, which can be used to calculate the critical constant c_{γ_2} for a given α . When $\gamma_2 = \phi/2$, it matches the confidence level of the two-sided hyperbolic band given by expression (2.11), whereas when $\gamma_2 = 0$, it matches the confidence level of the best two-sided three-segment band given by expression (2.9) when $c_{3,2,1} = c_{3,2,2}$ (see Liu and Hayter, 2007).

3.2.2 Area of confidence set corresponding to the outer-hyperbolic band

From the relationship between C_{γ_2} and R_{γ_2}, similar derivation as in Section 2.4.4 can be used to show that $Area(C_{\gamma_2}) = \hat{\sigma}^2|U|Area(R_{\gamma_2})$. Hence, from Figure 3.8, it is also clear that the region R_{γ_2} can be partitioned into

- four fans such as the one formed by the lines $U(1, A_1)^T$, $U(1, A)^T$ and the boundary of the region R_{γ_2}.
- four small right-angled triangles such as the one formed by $U(1, a_1)^T$, the boundary of the region R_{γ_2} and the bisector of the angle ϕ.
- four big right-angled triangles such as the one formed by $U(1, a)^T$, the boundary of the region R_{γ_2} and the line joining the two vertices that do not lie on the bisector of the angle ϕ.

Therefore, the area of the region R_{γ_2} can be calculated as the sum of four times the area of one of the sectors, four times the area of one of the small right-angled triangles and four times the area of one of the big right-angled triangles.

The area of one fan is $\frac{1}{2}c_{\gamma_2}^2\gamma_2$.

The area of one small right-angled triangle is $\frac{1}{2}c_{\gamma_2}^2 \tan(\frac{\phi}{2} - \gamma_2)$.

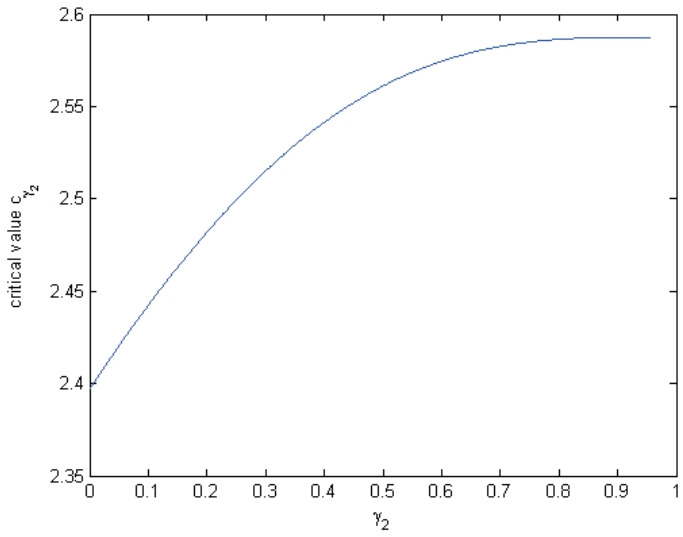
The area of one big right-angled triangle is $\frac{1}{2}c_{\gamma_2}^2 \cot(\frac{\phi}{2})$.

Therefore, the area of the confidence region R_{γ_2} is given by

$$Area(R_{\gamma_2}) = 2c_{\gamma_2}^2\gamma_2 + 2c_{\gamma_2}^2 \tan(\frac{\phi}{2} - \gamma_2) + 2c_{\gamma_2}^2 \cot(\frac{\phi}{2}).$$

When $\gamma_2 = 0$, $Area(R_{\gamma_2}) = 4c_{\gamma_2}^2/\sin\phi$, which is the area of the region R_3 for the best three-segment band. When $\gamma_2 = \phi/2$, $Area(R_{\gamma_2}) = c_{\gamma_2}^2\left[\phi + 2\cot(\frac{\phi}{2})\right]$, which is the area

Figure 3.10: Plot of c_{γ_2} against γ_2



of the region R_h for the hyperbolic band. Hence,

$$Area(C_{\gamma_2}) = \hat{\sigma}^2|U|\left[2c_{\gamma_2}^2\gamma_2 + 2c_{\gamma_2}^2\tan(\frac{\phi}{2}-\gamma_2) + 2c_{\gamma_2}^2\cot(\frac{\phi}{2})\right]. \tag{3.5}$$

3.2.3 Searching for the best outer-hyperbolic band

As in Section 3.1.3, the dataset on the desorption of carbon monoxide from Atkinson, Donev and Tobias (2007) given in Table 3.1 is used to illustrate the numerical search. The regression line b_0+b_1x is bounded over the range of interest $x \in (a, A) = (0, 2.5)$ by using a $1-\alpha = 0.95$ level simultaneous confidence band and the value of ϕ can be calculated using equation (2.8) to be 1.9167 *rad*. For each value of $\gamma_2 \in [0, \phi/2]$, the corresponding critical constant c_{γ_2} of the outer-hyperbolic band is computed using expression (3.4). The plot of c_{γ_2} against γ_2 is plotted in Figure 3.10. Then, the corresponding area of confidence set is calculated from expression (3.5) and depicted in Figure 3.11 from which the $\gamma_2 \in [0, \phi/2]$ that gives the MACS, i.e. the best outer-hyperbolic band, can be identified. Specifically, the area of the confidence set corresponding to the best outer-hyperbolic band is 0.07117 $units^2$, the optimal $\gamma_2 \in [0, \phi/2]$ is given by 0.6200 *rad* and the critical value $c_{\gamma_2} = 2.5765$. Using equation (3.1) and this optimal value of γ_2, the range (a_1, A_1) is found to be $(1.0421, 1.6559)$. The best outer-hyperbolic band is shown in Figure 3.12.

3.2.4 Comparisons

For given ϕ, ν and α, we can search numerically the best band in the family of $1-\alpha$ level outer-hyperbolic bands by finding the angle $\gamma_2 \in [0, \phi/2]$ that minimizes $Area(C_{\gamma_2})$ under the constraint $P\{\boldsymbol{T} \in R_{\gamma_2}\} = 1-\alpha$. We can also compare this best outer-hyperbolic band with the best three-segment band, with the hyperbolic band and with the best

Figure 3.11: Plot of $Area(C_{\gamma_2})$ against γ_2

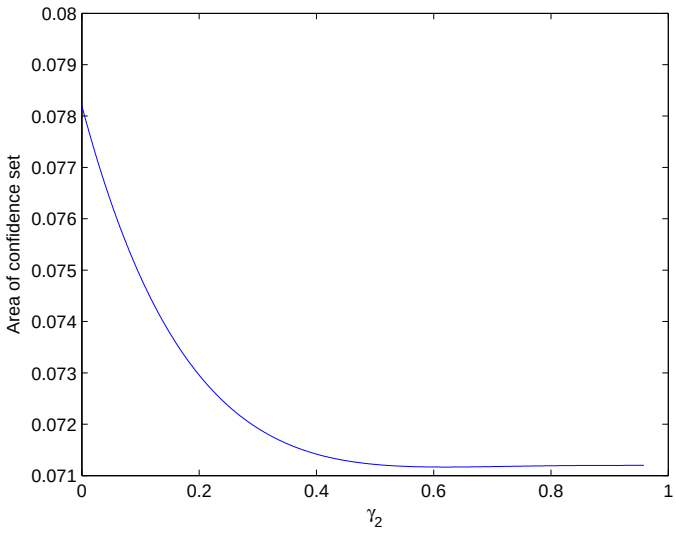


Figure 3.12: The best 0.95 level outer-hyperbolic band for the desorption CO dataset

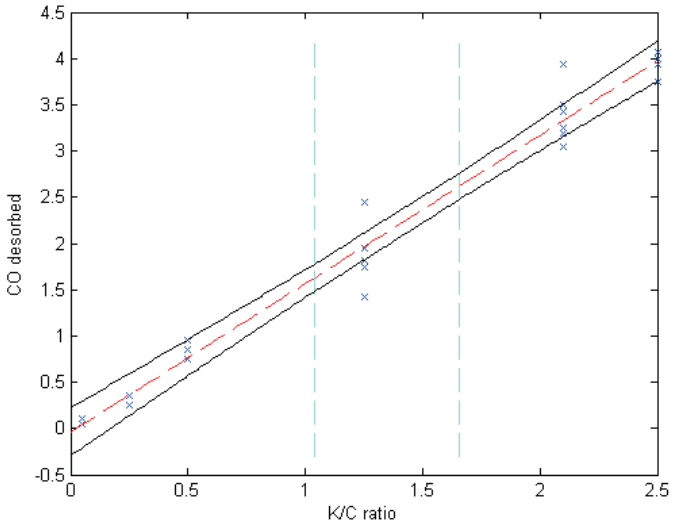
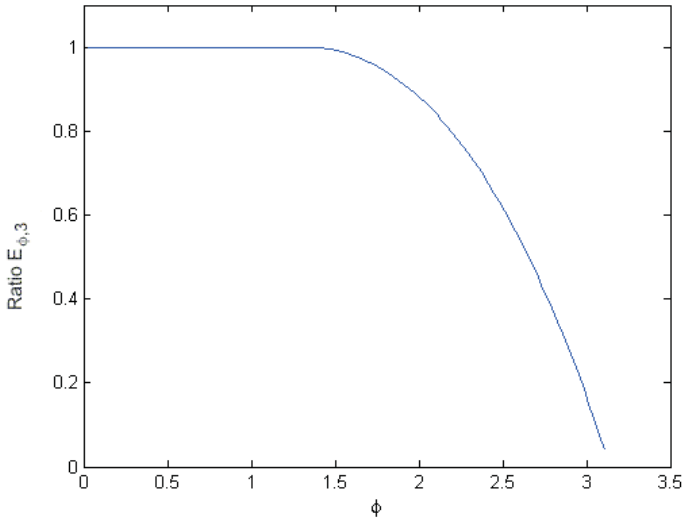


Figure 3.13: Plot of $E_{\phi,3}$ against ϕ



inner-hyperbolic band by looking at

$$E_{\phi,3} = \frac{Area(C_O^*)}{Area(C_3)} \quad , \quad E_{\phi,h} = \frac{Area(C_O^*)}{Area(C_h)} \quad \text{and} \quad E_{I,O} = \frac{Area(C_I^*)}{Area(C_O^*)}$$

where C_O^* denotes the confidence set of the best outer-hyperbolic band. For $E_{\phi,3}$ and $E_{\phi,h}$, similar observations as in Section 3.1.4 are made from the numerical investigation. As a function of $\phi \in (0, \pi)$, $E_{\phi,3}$ is equal to one for $\phi \in (0, \phi^\star)$ where $\phi^\star$ is a value depending on ν and α. Then, $E_{\phi,3}$ strictly decreases to zero for $\phi \in (\phi^\star, \pi)$ as depicted in Figure 3.13. Table 3.2 provides the value of $\phi^\star$ for some combinations of ν and α. From this, it can be concluded that the best outer-hyperbolic band is actually given by the best three-segment band for $\phi \in (0, \phi^\star)$, but a more efficient band than the best three-segment band can be found for $\phi \in (\phi^\star, \pi)$.

The function $E_{\phi,h}$ first strictly decreases from one and then strictly increases and approaches one over $\phi \in (0, \pi)$. The minimum value of $E_{\phi,h}$ is again only marginally smaller than one as shown in Figure 3.14. From this, it can be concluded that the best outer-hyperbolic band is always more efficient than the hyperbolic band under the MACS criterion, but the gain in efficiency of the best outer-hyperbolic band over the hyperbolic band is small.

Finally, the function $E_{I,O}$ is always no larger than one for $\phi \in (0, \pi)$ and is only very slightly less than one for ϕ near 1.5 as depicted in Figure 3.15. This implies that the best inner-hyperbolic band is at least as good as the best outer-hyperbolic band but is better by only a very small amount.

Table 3.2: Values of $\phi^{\star}$ in *rad* for combinations of $\alpha = 0.01,\ 0.05,\ 0.10$ and $\nu = 10,\ 30,\ \infty$

	$\alpha = 0.10$	$\alpha = 0.05$	$\alpha = 0.01$
$\nu = 10$	1.3383	1.3207	1.2893
$\nu = 30$	1.3405	1.31821	1.2733
$\nu = \infty$	1.3418	1.3173	1.2632

Figure 3.14: Plot of $E_{\phi,h}$ against ϕ

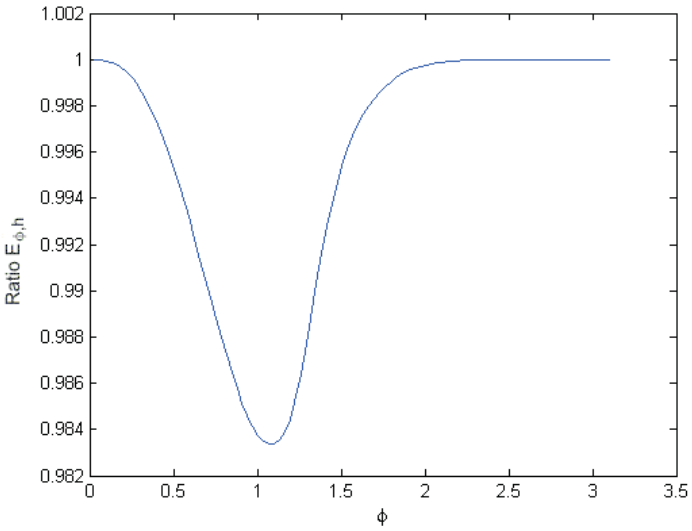
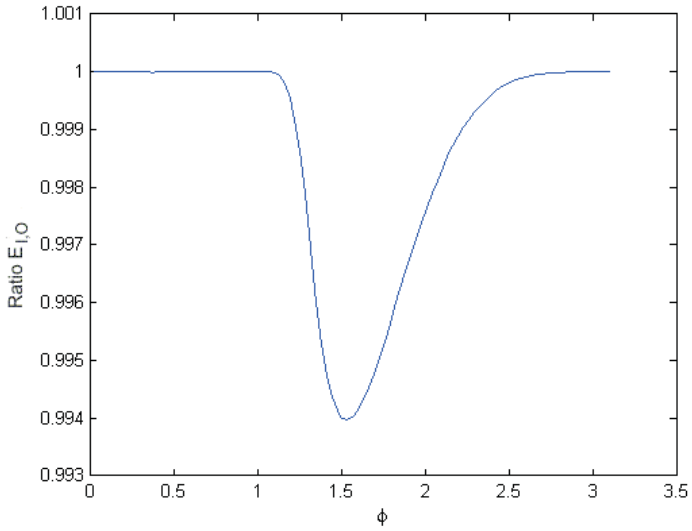


Figure 3.15: Plot of $E_{I,O}$ against ϕ



3.3 Concluding remarks on the inner-hyperbolic and outer-hyperbolic bands

Two new families of simultaneous confidence bands have been introduced and it is shown how the best confidence band in each family can be identified numerically. It is observed that the best inner-hyperbolic band is always no less efficient than the best outer-hyperbolic band and so the best inner-hyperbolic band is recommended.

The best inner-hyperbolic band is in fact given by the best three-segment band when $0 < \phi \leq \phi^*$ for some $\phi^* \in (0, \pi)$ depending on ν and α. But for $\phi^* < \phi < \pi$, the best inner-hyperbolic band can be much more efficient than the best three-segment band.

The best inner-hyperbolic band is always more efficient than the hyperbolic band, but only by a small amount. If one wants to avoid the burden of numerical search to find the best inner-hyperbolic band, then the hyperbolic band can be recommended with only a small loss of efficiency.

This concludes the work done on the construction and comparison of confidence bands in simple linear regression. In the next chapters, we turn our attention to the construction and comparison of confidence bands in multiple linear regression.

Chapter 4

Exact simultaneous confidence bands in multiple linear regression with predictor variables constrained in an ellipsoidal region

4.1 The ellipsoidal covariate region

Chapter 2 reviewed the construction of simultaneous confidence bands for simple linear regression, when $k = 1$. When $k > 1$, there are at least two predictor variables in the model (1.1) and the region of interest χ may assume various forms. The first part of this chapter reviews the construction of simultaneous confidence bands for the regression function

$$\boldsymbol{x}^T\boldsymbol{b} = b_0 + b_1x_1 + b_2x_2 + \ldots + b_kx_k$$

on an ellipsoidal region of interest χ_E. The linear regression model (1.2) is used, with the same assumptions and distributional results as in Chapter 1. Denote $X_{(1)}$ as the $n \times k$ matrix produced from the design matrix X by deleting the first column of 1's from X. Let $x_{.l} = \sum_{i=1}^{n} x_{il}$ be the mean of the observed values of the l^{th} predictor variable, where $1 \leq l \leq k$ and let $\bar{\boldsymbol{x}}_{(1)} = (x_{.1}, \ldots x_{.k})^T$. Then, let S be a $k \times k$ matrix given by

$$S \;=\; \frac{1}{n}\left(X_{(1)} - \mathbf{1}\bar{\boldsymbol{x}}_{(1)}^T\right)^T\left(X_{(1)} - \mathbf{1}\bar{\boldsymbol{x}}_{(1)}^T\right) \;=\; \frac{1}{n}\left(X_{(1)}^TX_{(1)} - n\bar{\boldsymbol{x}}_{(1)}\bar{\boldsymbol{x}}_{(1)}^T\right)$$

where $\mathbf{1}$ is an n-vector of 1's. Note that S is the sample variance-covariance matrix of the k predictor variables. Hence, the region χ_E is defined by

$$\chi_E = \left\{\boldsymbol{x}_{(1)} :\; \left(\boldsymbol{x}_{(1)} - \bar{\boldsymbol{x}}_{(1)}\right)^T S^{-1}\left(\boldsymbol{x}_{(1)} - \bar{\boldsymbol{x}}_{(1)}\right) \leq r^2\right\} \tag{4.1}$$

where $\boldsymbol{x}_{(1)} = (x_1, \ldots, x_k)^T$ and $r > 0$ is a constant that determines the size of χ_E. The region is centered at $\bar{\boldsymbol{x}}_{(1)}$ and its volume is given by

$$\begin{aligned}
& \int_{\chi_E} 1 \, d\boldsymbol{x}_{(1)} \\
= & \int_{\left(\boldsymbol{x}_{(1)}-\bar{\boldsymbol{x}}_{(1)}\right)^T S^{-1}\left(\boldsymbol{x}_{(1)}-\bar{\boldsymbol{x}}_{(1)}\right)\leq r^2} 1 \, d\boldsymbol{x}_{(1)} \\
= & \int_{\boldsymbol{\omega}_m^T\boldsymbol{\omega}_m \leq r^2} |S^{\frac{1}{2}}| \, d\boldsymbol{\omega}_m \qquad \left(\text{where } \boldsymbol{\omega}_m = S^{\frac{1}{2}}\left(\boldsymbol{x}_{(1)} - \bar{\boldsymbol{x}}_{(1)}\right)\right) \\
= & |S^{\frac{1}{2}}| \int_{\boldsymbol{\omega}_m^T\boldsymbol{\omega}_m \leq r^2} 1 \, d\boldsymbol{\omega}_m
\end{aligned}$$

where $\int_{\boldsymbol{\omega}_m^T\boldsymbol{\omega}_m \leq r^2} 1 \, d\boldsymbol{\omega}_m$ is a $(k-1)$-sphere with radius r. Therefore, χ_E is a $(k-1)$-sphere transformed linearly by $S^{\frac{1}{2}}$ to result into an ellipsoid.

Given that

$$(X^TX)^{-1} = \begin{pmatrix} \frac{1}{n} + \bar{\boldsymbol{x}}S^{-1}\bar{\boldsymbol{x}}' & -\bar{\boldsymbol{x}}S^{-1} \\ -S^{-1}\bar{\boldsymbol{x}}' & S^{-1} \end{pmatrix}$$

and that

$$\boldsymbol{x}^T(X^TX)^{-1}\boldsymbol{x} = \frac{1}{n}\left[1 + \left(\boldsymbol{x}_{(1)} - \bar{\boldsymbol{x}}_{(1)}\right)^T S^{-1}\left(\boldsymbol{x}_{(1)} - \bar{\boldsymbol{x}}_{(1)}\right)\right],$$

χ_E can also be expressed as

$$\chi_E = \left\{\boldsymbol{x}_{(1)} : \ \boldsymbol{x}^T(X^TX)^{-1}\boldsymbol{x} \leq \frac{1+r^2}{n}\right\}. \tag{4.2}$$

Since

$$Var(\boldsymbol{x}^T\boldsymbol{b}) = \frac{\sigma^2}{n}\left[1 + \left(\boldsymbol{x}_{(1)} - \bar{\boldsymbol{x}}_{(1)}\right)^T S^{-1}\left(\boldsymbol{x}_{(1)} - \bar{\boldsymbol{x}}_{(1)}\right)\right],$$

$Var(\boldsymbol{x}^T\boldsymbol{b}) = \frac{\sigma^2}{n}(1+r^2)$ for all the points $\boldsymbol{x}_{(1)}$ on the surface of the ellipsoidal region χ_E. Hence, all the points $\boldsymbol{x}_{(1)}$ on the surface of χ_E can be regarded as of equal "distance" in terms of $Var(\boldsymbol{x}^T\boldsymbol{b})$ from the center $\bar{\boldsymbol{x}}_{(1)}$. Therefore, the value of r^2 can be considered as the "range" for the region of interest χ_E.

Construction of confidence bands over regions like χ_E has been considered by Halperin and Guirian (1968), Bohrer (1973), Casella and Strawderman (1980) and Seppanen and Uusipaikka (1992) among others. Recently, Liu and Lin (2008) provided detailed construction of exact hyperbolic confidence bands over χ_E, while Liu *et al.* (2009) contained details on the construction of exact constant width bands over χ_E. In this chapter, the construction of two-sided hyperbolic and constant-width bands over χ_E are reviewed and a family of confidence bands, called the *inner-hyperbolic* bands, which include the hyperbolic and constant-width bands as special cases is introduced. The optimal confidence band within the family under the Minimum Volume Confidence Set (MVCS) criterion of Liu and Hayter (2007) and Liu *et al.* (2009) is found numerically and compared with the hyperbolic and constant width bands.

4.2 Preliminaries

As in Liu and Lin (2008), let $\boldsymbol{z} = \sqrt{n}(1, \bar{\boldsymbol{x}}_{(1)})^T$ and let the $(k+1) \times k$ matrix Z satisfy $(\boldsymbol{z}, Z)^T(X^TX)^{-1}(\boldsymbol{z}, Z) = I_{k+1}$. Then, it follows that $\boldsymbol{T} = (\boldsymbol{z}, Z)^{-1}(X^TX)(\hat{\boldsymbol{b}} - \boldsymbol{b})/\hat{\sigma}$ is a

standard t random vector of $k+1$ dimensions with $\nu = n-k-1$ degrees of freedom, denoted as $\boldsymbol{T} \sim t_{k+1,\nu}$. Moreover, let $\boldsymbol{w} = (\boldsymbol{z}, Z)^T (X^T X)^{-1} \boldsymbol{x} = (w_1, \boldsymbol{w}_{(1)})^T$, where $\boldsymbol{w}_{(1)} = (w_2, w_3, \ldots, w_{k+1})^T$, so that $w_1 = \boldsymbol{z}^T (X^T X)^{-1} \boldsymbol{x} = \frac{1}{\sqrt{n}}$, $\boldsymbol{w}_{(1)} = Z^T (X^T X)^{-1} \boldsymbol{x}$ and $\boldsymbol{w}^T \boldsymbol{w} = \boldsymbol{x} (X^T X)^{-1} \boldsymbol{x} = \|\boldsymbol{w}\|^2$. Then, from equation (4.2), all the possible values of $\boldsymbol{w}_{(1)}$, determined from $\boldsymbol{w} = (\boldsymbol{z}, Z)^T (X^T X)^{-1} \boldsymbol{x}$ when $\boldsymbol{x}_{(1)}$ varies over the region χ_E, form the set

$$W_E = \left\{ \boldsymbol{w} : \ w_1 = \frac{1}{\sqrt{n}}, \ \|\boldsymbol{w}\|^2 \leq \frac{1+r^2}{n} \right\}.$$

The hyperbolic band over χ_E has the form

$$\boldsymbol{x}^T \boldsymbol{b} \in \boldsymbol{x}^T \hat{\boldsymbol{b}} \pm c_h \hat{\sigma} \sqrt{\boldsymbol{x}^T (X^T X)^{-1} \boldsymbol{x}} \ \ \forall \boldsymbol{x}_{(1)} = (x_1, \ldots, x_k)^T \in \chi_E,$$

and its confidence level can be expressed as

$$\begin{aligned}
1-\alpha &= P\left\{ \sup_{\boldsymbol{x}_{(1)} \in \chi_E} \frac{|\boldsymbol{x}^T (\hat{\boldsymbol{b}} - \boldsymbol{b})|}{\hat{\sigma} \sqrt{\boldsymbol{x}^T (X^T X)^{-1} \boldsymbol{x}}} \leq c_h \right\} \\
&= P\left\{ \sup_{\boldsymbol{x}_{(1)} \in \chi_E} \frac{\left| \{(\boldsymbol{z}, Z)^T (X^T X)^{-1} \boldsymbol{x}\}^T \{(\boldsymbol{z}, Z)^{-1} (X^T X)(\hat{\boldsymbol{b}} - \boldsymbol{b})/\hat{\sigma}\} \right|}{\sqrt{\{(\boldsymbol{z}, Z)^T (X^T X)^{-1} \boldsymbol{x}\}^T \{(\boldsymbol{z}, Z)^T (X^T X)^{-1} \boldsymbol{x}\}}} \leq c_h \right\} \\
&= P\left\{ \sup_{\boldsymbol{x}_{(1)} \in \chi_E} \frac{\left| \{(\boldsymbol{z}, Z)^T (X^T X)^{-1} \boldsymbol{x}\}^T \boldsymbol{T} \right|}{\|(\boldsymbol{z}, Z)^T (X^T X)^{-1} \boldsymbol{x}\|} \leq c_h \right\} \\
&= P\{\boldsymbol{T} \in V_h\}
\end{aligned}$$

where

$$V_h = \left\{ \boldsymbol{T} : \ \sup_{\boldsymbol{w} \in W_E} \left| \boldsymbol{w}^T \boldsymbol{T} \right| \ / \ \|\boldsymbol{w}\| \leq c_h \right\}.$$

The region V_h is depicted in Figure 4.1, where the angle ϕ is given by

$$\phi = \cos^{-1}\left(\frac{1}{\sqrt{1+r^2}} \right) \ \in \ (0, \frac{\pi}{2}). \tag{4.3}$$

The constant width band over χ_E has the form

$$\boldsymbol{x}^T \boldsymbol{b} \in \boldsymbol{x}^T \hat{\boldsymbol{b}} \pm c_c \hat{\sigma} \sqrt{(1+r^2)/n} \ \ \forall \boldsymbol{x}_{(1)} = (x_1, \ldots, x_k)^T \in \chi_E,$$

and its confidence level can be expressed as

$$\begin{aligned}
1-\alpha &= P\left\{ \sup_{\boldsymbol{x}_{(1)} \in \chi_E} |\boldsymbol{x}^T (\hat{\boldsymbol{b}} - \boldsymbol{b})|/\hat{\sigma} \leq c_c \sqrt{(1+r^2)/n} \right\} \\
&= P\left\{ \sup_{\boldsymbol{x}_{(1)} \in \chi_E} \left| \{(\boldsymbol{z}, Z)^T (X^T X)^{-1} \boldsymbol{x}\}^T \{(\boldsymbol{z}, Z)^{-1} (X^T X)(\hat{\boldsymbol{b}} - \boldsymbol{b})/\hat{\sigma}\} \right| \leq c_c \sqrt{(1+r^2)/n} \right\} \\
&= P\left\{ \sup_{\boldsymbol{x}_{(1)} \in \chi_E} \left| \{(\boldsymbol{z}, Z)^T (X^T X)^{-1} \boldsymbol{x}\}^T \boldsymbol{T} \right| \leq c_c \sqrt{(1+r^2)/n} \right\} \\
&= P\{\boldsymbol{T} \in V_c\}
\end{aligned}$$

where

$$V_c = \left\{ \boldsymbol{T} : \ \sup_{\boldsymbol{w} \in W_E} \left| \boldsymbol{w}^T \boldsymbol{T} \right| \leq c_c \sqrt{(1+r^2)/n} \right\}.$$

Figure 4.1: Cross-section of V_h in the direction of t_1

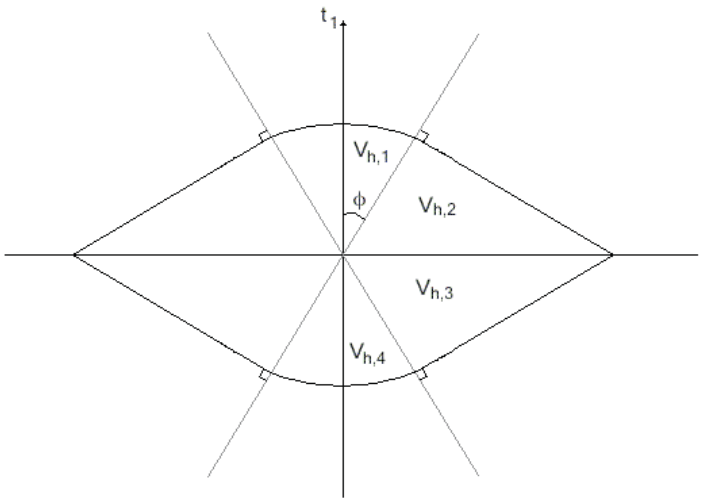
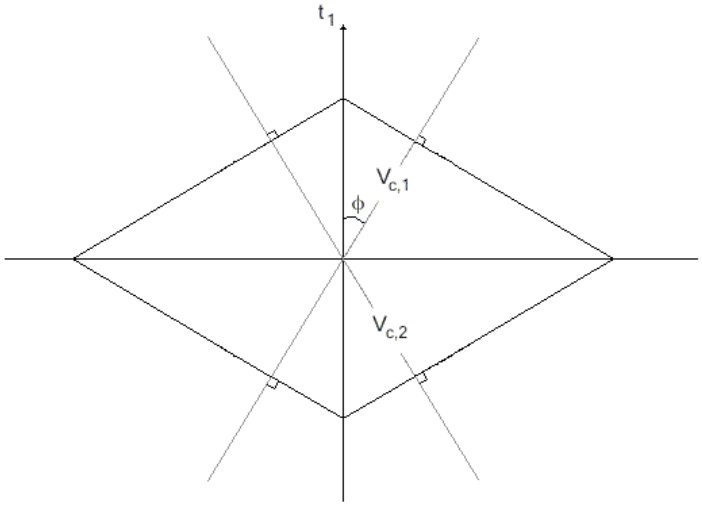


Figure 4.2: Cross-section of V_c in the direction of t_1



The region V_c is depicted in Figure 4.2.

Polar (hyperspherical) coordinates are used for the construction of the confidence bands in this paper, as in Liu and Lin (2008) and Liu *et al.* (2009), as well as the calculation of the volumes of the confidence sets. The polar coordinates $(R_{\boldsymbol{T}}, \theta_{\boldsymbol{T}1}, \ldots, \theta_{\boldsymbol{T}k})^T$ of the $(k+1)$-dimensional vector $\boldsymbol{T} = (t_1, \ldots, t_{k+1})^T$ are defined by

$$\begin{aligned} t_1 &= R_{\boldsymbol{T}} \cos\theta_{\boldsymbol{T}1} \\ t_2 &= R_{\boldsymbol{T}} \sin\theta_{\boldsymbol{T}1} \cos\theta_{\boldsymbol{T}2} \\ t_3 &= R_{\boldsymbol{T}} \sin\theta_{\boldsymbol{T}1} \sin\theta_{\boldsymbol{T}2} \cos\theta_{\boldsymbol{T}3} \\ &\vdots \quad\quad \vdots \\ t_k &= R_{\boldsymbol{T}} \sin\theta_{\boldsymbol{T}1} \sin\theta_{\boldsymbol{T}2} \ldots \sin\theta_{\boldsymbol{T}k-1} \cos\theta_{\boldsymbol{T}k} \\ t_{k+1} &= R_{\boldsymbol{T}} \sin\theta_{\boldsymbol{T}1} \sin\theta_{\boldsymbol{T}2} \ldots \sin\theta_{\boldsymbol{T}k-1} \sin\theta_{\boldsymbol{T}k} \end{aligned}$$

where

$$\begin{aligned} 0 &\leq \theta_{\boldsymbol{T}1} \leq \pi \\ 0 &\leq \theta_{\boldsymbol{T}2} \leq \pi \\ &\vdots \quad\quad \vdots \\ 0 &\leq \theta_{\boldsymbol{T}k-1} \leq \pi \\ 0 &\leq \theta_{\boldsymbol{T}k} \leq 2\pi \\ R_{\boldsymbol{T}} &\geq 0. \end{aligned}$$

When $\boldsymbol{T} \sim t_{k+1,\nu}$, the Jacobian of the transformation is

$$|J| = R_{\boldsymbol{T}}^k \sin^{k-1}\theta_{\boldsymbol{T}1} \sin^{k-2}\theta_{\boldsymbol{T}2} \ldots \sin\theta_{\boldsymbol{T}k-1}. \tag{4.4}$$

Its polar coordinates are independent random variables. In particular, the marginal density of $\theta_{\boldsymbol{T}1}$ is given by

$$f(\theta) = g \sin^{k-1}\theta \quad 0 \leq \theta \leq \pi$$

where g is the normalizing constant given by $g = \frac{1}{\int_0^\pi \sin^{k-1}\theta\, d\theta}$ and the marginal distribution of $R_{\boldsymbol{T}}$ is given by

$$R_{\boldsymbol{T}} \sim \sqrt{(k+1)F_{(k+1),\nu}}$$

where $F_{(k+1),\nu}$ denotes an F random variable that has $(k+1)$ and ν degrees of freedom.

Let $v(R)$ denote the volume of a set $R \subset \Re^{(k+1)}$ and let $B_{k+1}(p)$ denote the ball of radius p in $\Re^{(k+1)}$. Using the Jacobian of the transformation from cartesian to polar coordinates in (4.4), the volume of the ball, $v\left(B_{k+1}(p)\right)$, can be expressed as

$$v\left(B_{k+1}(p)\right) = \int_{R=0}^{p} \int_{\theta_1=0}^{\pi} \int_{\theta_2=0}^{\pi} \cdots \int_{\theta_{k-1}=0}^{\pi} \int_{\theta_k=0}^{2\pi} |J| dR d\theta_1 d\theta_2 \ldots d\theta_k = \frac{\pi^{\frac{k+1}{2}} p^{k+1}}{\Gamma\left[\frac{k+1}{2}+1\right]}.$$

The expression for the volume of a ball of radius p is used in the following sections of this chapter to derive expressions for volumes of confidence sets corresponding to simultaneous confidence bands.

4.3 Two-sided hyperbolic band over χ_E

4.3.1 Confidence level

The confidence level of the band is given by $P\{\boldsymbol{T} \in V_h\}$ and the region V_h can be partitioned into the regions $V_{h,1}$, $V_{h,2}$, $V_{h,3}$ and $V_{h,4}$ as depicted in Figure 4.1, where

$$\begin{aligned}
V_{h,1} &= \{\boldsymbol{T} :\ 0 < \theta_{\boldsymbol{T}1} \leq \phi,\ R_{\boldsymbol{T}} \leq c_h\}\,, \\
V_{h,2} &= \left\{\boldsymbol{T} :\ \phi < \theta_{\boldsymbol{T}1} \leq \frac{\pi}{2},\ R_{\boldsymbol{T}} \cos(\theta_{\boldsymbol{T}1} - \phi) \leq c_h\right\}, \\
V_{h,3} &= \left\{\boldsymbol{T} :\ \frac{\pi}{2} < \theta_{\boldsymbol{T}1} \leq \pi - \phi,\ R_{\boldsymbol{T}} \cos(\pi - \theta_{\boldsymbol{T}1} - \phi) \leq c_h\right\}, \\
V_{h,4} &= \{\boldsymbol{T} :\ \pi - \phi < \theta_{\boldsymbol{T}1} < \pi,\ R_{\boldsymbol{T}} \leq c_h\}\,.
\end{aligned}$$

Due to the rotational invariance of the probability distribution of $\boldsymbol{T}$, the probability of $\boldsymbol{T}$ in $V_{h,1}$ is given by

$$\begin{aligned}
P\{\boldsymbol{T} \in V_{h,1}\} &= P\{\boldsymbol{T} \in V_{h,4}\} \\
&= \int_0^{\phi} g \sin^{k-1}\theta \, d\theta \ .\ P\{R_{\boldsymbol{T}} \leq c_h\} \\
&= \int_0^{\phi} g \sin^{k-1}\theta \, d\theta \ .\ P\{(k+1)F_{(k+1),\nu} \leq c_h^2\} \\
&= \int_0^{\phi} g \sin^{k-1}\theta \, d\theta \ .\ F_{(k+1),\nu}\left(\frac{c_h^2}{k+1}\right).
\end{aligned}$$

Similarly,

$$\begin{aligned}
P\{\boldsymbol{T} \in V_{h,2}\} &= P\{\boldsymbol{T} \in V_{h,3}\} \\
&= \int_{\phi}^{\frac{\pi}{2}} g \sin^{k-1}\theta \ .\ P\{R_{\boldsymbol{T}} \cos(\theta - \phi) \leq c_h\}\, d\theta \\
&= \int_0^{\frac{\pi}{2}-\phi} g \sin^{k-1}(\theta + \phi) \ .\ P\{R_{\boldsymbol{T}} \leq \frac{c_h}{\cos\theta}\}\, d\theta \\
&= \int_0^{\frac{\pi}{2}-\phi} g \sin^{k-1}(\theta + \phi) \ .\ F_{(k+1),\nu}\left(\frac{c_h^2}{(k+1)\cos^2\theta}\right) d\theta.
\end{aligned}$$

The confidence level of the two-sided hyperbolic band over χ_E is therefore given by

$$\begin{aligned}
1 - \alpha &= \int_0^{\phi} 2g \sin^{k-1}\theta \, d\theta \ .\ F_{(k+1),\nu}\left(\frac{c_h^2}{k+1}\right) \\
&\quad + \int_0^{\frac{\pi}{2}-\phi} 2g \sin^{k-1}(\theta + \phi) \ .\ F_{(k+1),\nu}\left(\frac{c_h^2}{(k+1)\cos^2\theta}\right) d\theta. \qquad (4.5)
\end{aligned}$$

4.3.2 Volume of confidence set

Using the partitioning $V_h = V_{h,1} + V_{h,2} + V_{h,3} + V_{h,4}$, we have $v(V_h) = v(V_{h,1}) + v(V_{h,2}) + v(V_{h,3}) + v(V_{h,4})$, with

$$\begin{aligned} & v(V_{h,1}) \;=\; v(V_{h,4}) \\ = \;& \int_{R=0}^{c_h} \int_{\theta_1=0}^{\phi} \int_{\theta_2=0}^{\pi} \cdots \int_{\theta_{k-1}=0}^{\pi} \int_{\theta_k=0}^{2\pi} |J| dR d\theta_1 d\theta_2 d\theta_k \\ = \;& \left\{ v\left(B_{k+1}(c_h)\right) \;/\; \int_0^{\pi} \sin^{k-1}\theta_1 \, d\theta_1 \right\} \;.\; \int_0^{\phi} \sin^{k-1}\theta_1 \, d\theta_1 \\ = \;& g \int_0^{\phi} \sin^{k-1}\theta_1 \, d\theta_1 \;.\; v\left(B_{k+1}(c_h)\right) \end{aligned}$$

and

$$\begin{aligned} & v(V_{h,2}) \;=\; v(V_{h,3}) \\ = \;& \iint_{R_{\boldsymbol{T}} \cos(\theta_1 - \phi) \le c_h} \int_{\theta_2=0}^{\pi} \cdots \int_{\theta_{k-1}=0}^{\pi} \int_{\theta_k=0}^{2\pi} |J| dR d\theta_1 d\theta_2 d\theta_k \\ & \qquad \phi < \theta_1 \le \tfrac{\pi}{2} \\ = \;& \left\{ v\left(B_{k+1}(c_h)\right) / \int_{R=0}^{c_h} \int_{\theta_1=0}^{\pi} R^k \sin^{k-1}\theta_1 \, dR d\theta_1 \right\} \\ \times \;& \iint_{R_{\boldsymbol{T}} \cos(\theta_1 - \phi) \le c_h} R^k \sin^{k-1}\theta_1 \, dR d\theta_1 \\ & \qquad \phi < \theta_1 \le \tfrac{\pi}{2} \\ = \;& g \int_{\theta_1=\theta}^{\frac{\pi}{2}} \frac{\sin^{k-1}\theta_1}{\cos^{k+1}(\theta_1 - \phi)} \, d\theta_1 \;.\; v\left(B_{k+1}(c_h)\right). \end{aligned}$$

Therefore, the volume of the confidence region V_h is given by

$$v(V_h) = 2g \frac{\pi^{\frac{k+1}{2}} c_h^{k+1}}{\Gamma\left[\frac{k+1}{2} + 1\right]} \left[\int_0^{\phi} \sin^{k-1}\theta_1 \, d\theta_1 + \int_{\theta_1=\theta}^{\frac{\pi}{2}} \frac{\sin^{k-1}\theta_1}{\cos^{k+1}(\theta_1 - \phi)} \, d\theta_1 \right].$$

The confidence set corresponding to the two-sided hyperbolic band over χ_E, C_h, has the form

$$C_h = \left\{ \boldsymbol{b} : \; (\boldsymbol{z}, Z)^{-1} (X^T X)(\hat{\boldsymbol{b}} - \boldsymbol{b}) / \hat{\sigma} \in V_h \right\}$$

which satisfies

$$P\{\boldsymbol{b} \in C_h\} = P\{\boldsymbol{T} \in V_h\} = 1 - \alpha$$

and can be expressed as a linear transformation of V_h:

$$C_h = \left\{ \boldsymbol{b} : \; \boldsymbol{b} \in \hat{\boldsymbol{b}} + \hat{\sigma} (X^T X)^{-1} (\boldsymbol{z}, Z) V_h \right\}.$$

Since

$$v(C_h) = |\hat{\sigma} (X^T X)^{-1} (\boldsymbol{z}, Z)| v(V_h) = \hat{\sigma}^{k+1} |(X^T X)^{-\frac{1}{2}}| v(V_h),$$

the volume of C_h is given by

$$2\hat{\sigma}^{k+1} |(X^T X)^{-\frac{1}{2}}| g \frac{\pi^{\frac{k+1}{2}} c_h^{k+1}}{\Gamma\left[\frac{k+1}{2} + 1\right]} \left[\int_0^{\phi} \sin^{k-1}\theta_1 \, d\theta_1 + \int_{\theta_1=\theta}^{\frac{\pi}{2}} \frac{\sin^{k-1}\theta_1}{\cos^{k+1}(\theta_1 - \phi)} \, d\theta_1 \right]. \qquad (4.6)$$

4.4 Two-sided constant width band over χ_E

4.4.1 Confidence level

The confidence level of the band is given by $P\{\boldsymbol{T} \in V_c\}$, where

$$V_c = \left\{ \boldsymbol{T} : \sup_{\boldsymbol{w} \in W_E} \left| \boldsymbol{w}^T \boldsymbol{T} \right| \leq c_c \sqrt{(1+r^2)/n} \right\}.$$

Note that

$$\begin{aligned} & \sup_{\boldsymbol{w} \in W_E} \left| \boldsymbol{w}^T \boldsymbol{T} \right| \\ = & \sup_{\boldsymbol{w} \in W_E} \left| \frac{t_1}{\sqrt{n}} \boldsymbol{w}_{(1)}^T \boldsymbol{T}_{(1)} \right| \\ = & \frac{|t_1|}{\sqrt{n}} + \sqrt{\frac{r^2}{n}} \|\boldsymbol{T}_{(1)}\| \end{aligned}$$

where $\boldsymbol{T} = (t_1, \boldsymbol{T}_{(1)})^T = (t_1, t_2, \ldots, t_{k+1})^T$. Then,

$$V_c = \left\{ \boldsymbol{T} : \frac{|t_1|}{\sqrt{n}} + \sqrt{\frac{r^2}{n}} \|\boldsymbol{T}_{(1)}\| \leq c_c \sqrt{(1+r^2)/n} \right\}.$$

In polar coordinates,

$$\begin{aligned} V_c &= \left\{ \boldsymbol{T} : \frac{|R_{\boldsymbol{T}} \cos\theta_{\boldsymbol{T}1}|}{\sqrt{n}} + \sqrt{\frac{r^2}{n}} |R_{\boldsymbol{T}} \cos\theta_{\boldsymbol{T}1}| \leq c_c \sqrt{(1+r^2)/n} \right\} \\ &= V_{c,1} + V_{c,2} \end{aligned}$$

as depicted in Figure 4.2. The regions $V_{c,1}$ and $V_{c,2}$ can be expressed as

$$V_{c,1} = \left\{ \boldsymbol{T} : \ 0 < \theta_{\boldsymbol{T}1} \leq \frac{\pi}{2}, \ R_{\boldsymbol{T}} \cos(\theta_{\boldsymbol{T}1} - \phi) \leq c_c \right\},$$

$$V_{c,2} = \left\{ \boldsymbol{T} : \ \frac{\pi}{2} < \theta_{\boldsymbol{T}1} < \pi, \ R_{\boldsymbol{T}} \cos(\pi - \theta_{\boldsymbol{T}1} - \phi) \leq c_c \right\}.$$

Since $P\{\boldsymbol{T} \in V_{c,1}\} = P\{\boldsymbol{T} \in V_{c,2}\}$,

$$\begin{aligned} & P\{\boldsymbol{T} \in V_c\} \\ = & \ 2P\{\boldsymbol{T} \in V_{c,1}\} \\ = & \ 2 \int_0^{\frac{\pi}{2}} g \sin^{k-1}\theta \ . \ P\{R_{\boldsymbol{T}} \cos(\theta - \phi) \leq c_c\} \, d\theta \\ = & \ 2 \int_0^{\frac{\pi}{2}} g \sin^{k-1}\theta \ . \ F_{(k+1),\nu}\left(\frac{c_c^2}{(k+1)\cos^2(\theta - \phi)} \right) d\theta. \end{aligned}$$

The confidence level of the two-sided constant-width band over χ_E is therefore given by

$$1 - \alpha = 2 \int_0^{\frac{\pi}{2}} g \sin^{k-1}\theta \ . \ F_{(k+1),\nu}\left(\frac{c_c^2}{(k+1)\cos^2(\theta - \phi)} \right) d\theta. \tag{4.7}$$

4.4.2 Volume of confidence set

Using the partitioning $V_c = V_{c,1} + V_{c,2}$, we have $v(V_c) = v(V_{c,1}) + v(V_{c,2})$, with

$$\begin{aligned} & v(V_{c,1}) \;=\; v(V_{c,2}) \\ = \quad & \iint_{R_{\boldsymbol{T}}\cos(\theta_1-\phi)\le c_c,\; 0<\theta_1\le\frac{\pi}{2}} \int_{\theta_2=0}^{\pi} \cdots \int_{\theta_{k-1}=0}^{\pi} \int_{\theta_k=0}^{2\pi} |J| dR d\theta_1 d\theta_2 d\theta_k \\ = \quad & \left\{ v\left(B_{k+1}(c_c)\right) / \int_{R=0}^{c_c}\int_{\theta_1=0}^{\pi} R^k \sin^{k-1}\theta_1 \, dR d\theta_1 \right\} \\ \times \quad & \iint_{R_{\boldsymbol{T}}\cos(\theta_1-\phi)\le c_c,\; 0<\theta_1\le\frac{\pi}{2}} R^k \sin^{k-1}\theta_1 \, dR d\theta_1 \\ = \quad & g\int_0^{\frac{\pi}{2}} \frac{\sin^{k-1}\theta_1}{\cos^{k+1}(\theta_1-\phi)}\, d\theta_1 \;.\; v\left(B_{k+1}(c_c)\right). \end{aligned}$$

Therefore, the volume of the confidence region V_c is given by

$$v(V_c) = 2g\frac{\pi^{\frac{k+1}{2}} c_c^{k+1}}{\Gamma\left[\frac{k+1}{2}+1\right]} \int_0^{\frac{\pi}{2}} \frac{\sin^{k-1}\theta_1}{\cos^{k+1}(\theta_1-\phi)}\, d\theta_1 \;.\; v\left(B_{k+1}(c_c)\right).$$

The confidence set corresponding to the two-sided constant-width band over χ_E, C_c, has the form

$$C_c = \left\{ \boldsymbol{b} :\; (\boldsymbol{z}, Z)^{-1}(X^T X)(\hat{\boldsymbol{b}} - \boldsymbol{b})/\hat{\sigma} \in V_c \right\}$$

which satisfies

$$P\{\boldsymbol{b} \in C_c\} = P\{\boldsymbol{T} \in V_c\} = 1 - \alpha$$

and can be expressed as a linear transformation of V_c:

$$C_c = \left\{ \boldsymbol{b} :\; \boldsymbol{b} \in \hat{\boldsymbol{b}} + \hat{\sigma}(X^T X)^{-1}(\boldsymbol{z}, Z) V_c \right\}.$$

Since

$$v(C_c) = |\hat{\sigma}(X^T X)^{-1}(\boldsymbol{z}, Z)| v(V_c) = \hat{\sigma}^{k+1} |(X^T X)^{-\frac{1}{2}}| v(V_c),$$

the expression for the volume of C_c is given by

$$2\hat{\sigma}^{k+1}|(X^T X)^{-\frac{1}{2}}| g\frac{\pi^{\frac{k+1}{2}} c_c^{k+1}}{\Gamma\left[\frac{k+1}{2}+1\right]} \int_0^{\frac{\pi}{2}} \frac{\sin^{k-1}\theta_1}{\cos^{k+1}(\theta_1-\phi)}\, d\theta_1. \tag{4.8}$$

4.5 Numerical example

The two-sided hyperbolic and constant-width bands constructed in this chapter can be used for linear regression models where $k = 1, 2$ or more than 2. When $k = 1$ or $k = 2$, graphical representation of the bands is possible. However, when $k > 2$, the only illustrations possible are cross-sections along certain planes. A portion of the Snee (1977) acetylene dataset, shown in Table 4.1, is used to illustrate the case when $k = 2$. The first two predictor variables are used, namely the reactor temperature x_1 and the ratio of H_2 to n-Heptane x_2 and the response variable is the conversion of n-Heptane to Acetylene y.

Table 4.1: Snee (1977) Acetylene dataset

Conversion of n-Heptane to Acetylene y (%)	Reactor temperature x_1 ($^\circ C$)	ratio of H_2 to n-Heptane x_2 (mole ratio)	Contact time x_3 (seconds)
49.0	1300	7.5	0.0120
50.2	1300	9.0	0.0120
50.5	1300	11.0	0.0115
48.5	1300	13.5	0.0130
47.5	1300	17.0	0.0135
44.5	1300	23.0	0.0120
28.0	1200	5.3	0.0400
31.5	1200	7.5	0.0380
34.5	1200	11.0	0.0320
35.0	1200	13.5	0.0260
38.0	1200	17.0	0.0340
38.5	1200	23.0	0.0410
15.0	1100	5.3	0.0840
17.0	1100	7.5	0.0980
20.5	1100	11.0	0.0920
29.5	1100	17.0	0.0860

There are sixteen observations in the dataset, $n = 16$, and the fitted regression model is given by

$$y = -130.69 + 0.134x_1 + 0.351x_2, \text{ with } \hat{\sigma} = 3.624 \text{ and } R^2 = 0.92$$

where R^2 is a popular exploratory measure of how well the model fits the observed data and can be interpreted as the proportion of the total variation in the response values that is explained by the systematic component $\boldsymbol{x}^T\boldsymbol{b}$ of the model. The ellipsoidal region χ_E is centered around the mean $\bar{\boldsymbol{x}}_{(1)} = (1212.5, 12.4)^T$ and its size increases as r is increased. We assume that we wish to bound the regression function $\boldsymbol{x}^T\boldsymbol{b}$ using a $1 - \alpha = 0.90$ level simultaneous confidence band over the ellipsoidal region χ_E with $r = 1.9$, so that the region of interest χ_E is given by

$$\chi_E = \left\{\boldsymbol{x}_{(1)} : \ \boldsymbol{x}^T(X^TX)^{-1}\boldsymbol{x} \leq 0.288125\right\}$$

as depicted in Figure 4.3 by the ellipse in the (x_1, x_2)-plane with $\phi = 1.0863$. Using expression (4.5), the critical constant c_h is found to be 2.7229. Hence, using expression (4.6), the volume of the confidence set corresponding to the hyperbolic band over χ_E is calculated to be 0.6419 $units^3$. The two-sided hyperbolic band over χ_E is illustrated in Figure 4.3. The hyperbolic shape of the band cannot be easily distinguished from the shape of a constant width band. Therefore, cross-sections of the band along the x_1-direction at $x_2 = x_{\cdot 2}$ and along the x_2-direction at $x_1 = x_{\cdot 1}$ are used to show the shape of the band.

Figure 4.3: Two-sided hyperbolic band over χ_E, Snee (1977) acetylene dataset

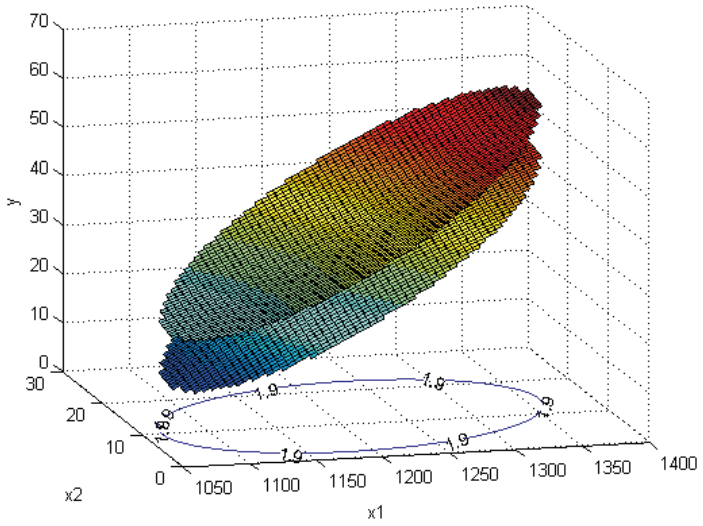


Figure 4.4 and Figure 4.5 show these respective cross-sections, where the hyperbolic form is clearer.

Similarly, using expression (4.7), the critical constant c_c is found to be 2.5981. Hence, using expression (4.8), the volume of the confidence set corresponding to the constant-width band over χ_E is found to be 0.7182 $units^3$. The band, when illustrated, appears fairly similar to the band shown in Figure 4.3, but the cross-sections of the the band along the x_1-direction at $x_2 = x_{\cdot 2}$ and along the x_2-direction at $x_1 = x_{\cdot 1}$, depicted in Figure 4.6 and Figure 4.7, clearly show the distinction from the hyperbolic shape.

Therefore, it can be deduced that for this example, the hyperbolic band is better than the constant width band under the MVCS criterion. In the next section, we introduce the family of inner-hyperbolic bands, first discussed in Chapter 3, in multiple linear regression over the region χ_E.

4.6 Family of inner-hyperbolic bands over χ_E

This family of confidence bands is defined in terms of regions V for $\boldsymbol{T} \sim t_{k+1,\nu}$, which are in turn used to construct confidence sets for $\boldsymbol{b}$. Each region V_γ is defined in terms of an angle $\gamma \in [0, \phi]$ in the following way. The region V_γ is given by

$$V_\gamma = V_{\gamma,1} + V_{\gamma,2} + V_{\gamma,3} + V_{\gamma,4}$$

Figure 4.4: Cross-section of the hyperbolic band along the x_1-direction at $x_2 = x_{\cdot 2}$

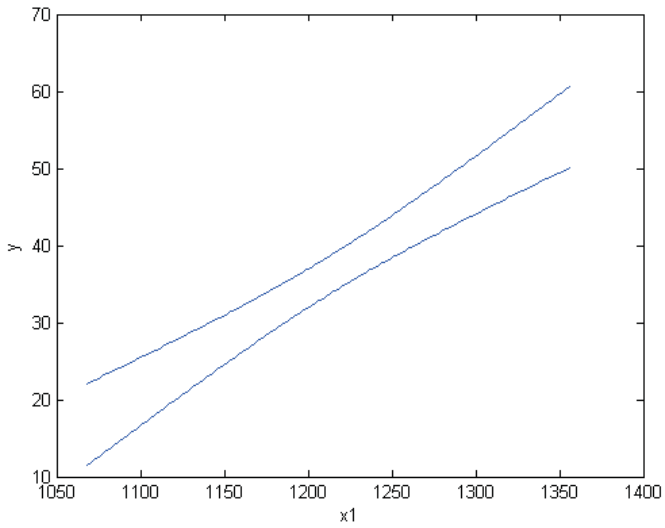


Figure 4.5: Cross-section of the hyperbolic band along the x_2-direction at $x_1 = x_{\cdot 1}$

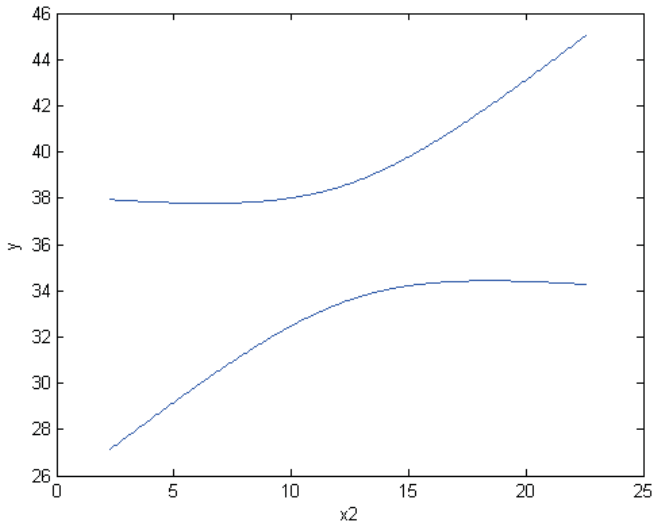


Figure 4.6: Cross-section of the constant width band along the x_1-direction at $x_2 = x_{\cdot 2}$

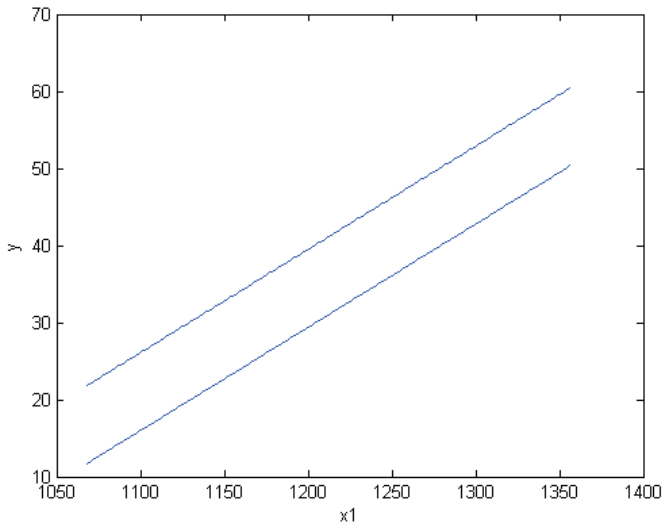


Figure 4.7: Cross-section of the constant width band along the x_2-direction at $x_1 = x_{\cdot 1}$

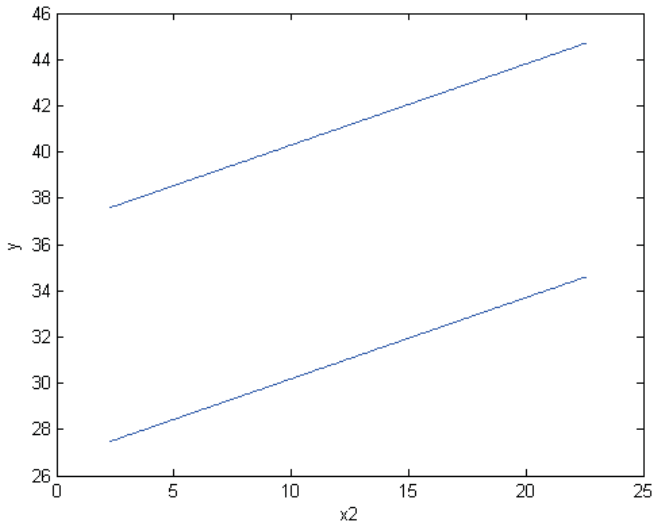
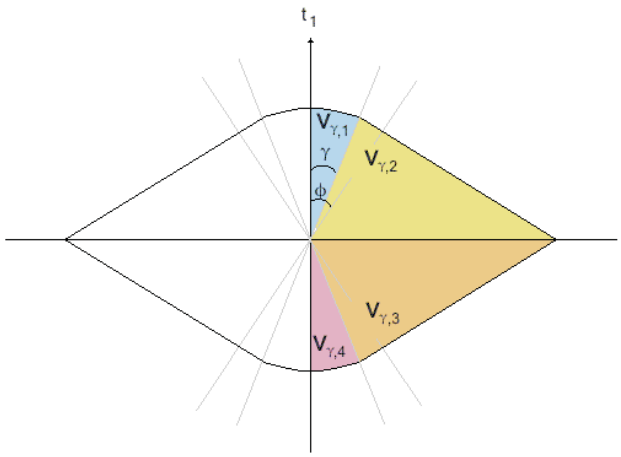


Figure 4.8: Cross-section of V_γ in the direction of t_1



where

$$V_{\gamma,1} = \{\boldsymbol{T} : \ 0 < \theta_{\boldsymbol{T}1} \leq \gamma, \ R_{\boldsymbol{T}} \cos(\phi - \gamma) \leq c_\gamma\},$$

$$V_{\gamma,2} = \left\{\boldsymbol{T} : \ \gamma < \theta_{\boldsymbol{T}1} \leq \frac{\pi}{2}, \ R_{\boldsymbol{T}} \cos(\theta_{\boldsymbol{T}1} - \phi) \leq c_\gamma\right\},$$

$$V_{\gamma,3} = \left\{\boldsymbol{T} : \ \frac{\pi}{2} < \theta_{\boldsymbol{T}1} \leq \pi - \gamma, \ R_{\boldsymbol{T}} \cos(\pi - \theta_{\boldsymbol{T}1} - \phi) \leq c_\gamma\right\},$$

$$V_{\gamma,4} = \{\boldsymbol{T} : \ \pi - \gamma < \theta_{\boldsymbol{T}1} < \pi, \ R_{\boldsymbol{T}} \cos(\phi - \gamma) \leq c_\gamma\}.$$

The region V_γ is depicted in Figure 4.8. The value of the critical constant c_γ is chosen so that $P\{\boldsymbol{T} \in V_\gamma\} = 1 - \alpha$ and hence c_γ depends on γ, α, k and ν and denoted by $c_\gamma = c_\gamma(\gamma, \alpha, k, \nu)$. It is clear that when $\gamma = \phi$, the region V_γ is simply the V_h depicted in Figure 4.1 corresponding to the hyperbolic band and when $\gamma = 0$, the region V_γ is simply the V_c depicted in Figure 4.2 corresponding to the constant width band. The confidence set for $\boldsymbol{b}$ corresponding to V_γ is given by

$$C_\gamma = \left\{\boldsymbol{b} : \ (\boldsymbol{z}, Z)^{-1}(X^T X)(\hat{\boldsymbol{b}} - \boldsymbol{b})/\hat{\sigma} \in V_\gamma\right\},$$

which has an exact confidence level $1 - \alpha$.

Now we give the confidence band that corresponds to the confidence set C_γ. For given $\gamma \in [0, \phi]$, a value r_γ $(0 \leq r_\gamma \leq r)$ can be solved uniquely from the relation

$$\gamma = \cos^{-1}\left(\frac{1}{\sqrt{1 + r_\gamma^2}}\right). \tag{4.9}$$

By comparing this with the equation in (4.3), it is clear that $r_\gamma \leq r$. Now, denote $\chi_{\gamma,E}$ as a covariate region given by

$$\begin{aligned}\chi_{\gamma,E} &= \left\{\boldsymbol{x}_{(1)} :\ \left(\boldsymbol{x}_{(1)} - \bar{\boldsymbol{x}}_{(1)}\right)^T S^{-1} \left(\boldsymbol{x}_{(1)} - \bar{\boldsymbol{x}}_{(1)}\right) \leq r_\gamma^2\right\} \\ &= \left\{\boldsymbol{x}_{(1)} :\ \boldsymbol{x}^T (X^T X)^{-1} \boldsymbol{x} \leq \frac{1 + r_\gamma^2}{n}\right\}.\end{aligned} \tag{4.10}$$

The region $\chi_{\gamma,E}$ is of similar shape as but smaller in size than the region χ_E. Thus, $\chi_{\gamma,E}$ can be regarded as the "inner-region" of the region of interest χ_E. Define χ_{oE} as the region within χ_E and outside of $\chi_{\gamma,E}$, given by

$$\chi_{oE} = \left\{\boldsymbol{x}_{(1)} :\ \frac{1 + r_\gamma^2}{n} \leq \boldsymbol{x}^T (X^T X)^{-1} \boldsymbol{x} \leq \frac{1 + r^2}{n}\right\}. \tag{4.11}$$

Now, the confidence band that corresponds to the confidence set C_γ can be shown to be given by

$$\boldsymbol{x}^T \boldsymbol{b} \in \boldsymbol{x}^T \hat{\boldsymbol{b}} \pm \hat{\sigma} H_\gamma(\boldsymbol{x}) \ \ \forall \boldsymbol{x}_{(1)} = (x_1, \ldots, x_k)^T \in \chi_E,$$

where

$$H_\gamma(\boldsymbol{x}) = \begin{cases} \frac{c_\gamma}{\cos(\phi - \gamma)} \sqrt{\boldsymbol{x}^T (X^T X)^{-1} \boldsymbol{x}} & \text{for } \forall \boldsymbol{x}_{(1)} \in \chi_{\gamma,E} \\ \frac{1}{(r - r_\gamma)} \left(r - \sqrt{n \boldsymbol{x}^T (X^T X)^{-1} \boldsymbol{x} - 1}\right) \frac{c_\gamma}{\cos(\phi - \gamma)} \sqrt{\frac{1 + r_\gamma^2}{n}} & \\ + \frac{1}{(r - r_\gamma)} \left(\sqrt{n \boldsymbol{x}^T (X^T X)^{-1} \boldsymbol{x} - 1} - r_\gamma\right) c_\gamma \sqrt{\frac{1 + r^2}{n}} & \text{for } \forall \boldsymbol{x}_{(1)} \in \chi_{oE} \end{cases}.$$

So for each $\gamma \in [0, \phi]$, we are able to define an exact $1 - \alpha$ level confidence band over χ_E. When γ varies over the interval $[0, \phi]$, we have a family of exact confidence bands over χ_E. In particular, $\gamma = 0$ corresponds to the constant width band and $\gamma = \phi$ corresponds to the hyperbolic band.

4.6.1 Confidence level

Now we discuss the computation of the critical constant $c_\gamma = c_\gamma(\gamma, \alpha, k, \nu)$. From the construction above, c_γ is determined from $P\{\boldsymbol{T} \in V_\gamma\} = 1 - \alpha$. Hence we need to express $P\{\boldsymbol{T} \in V_\gamma\} = P\{\boldsymbol{T} \in V_{\gamma,1}\} + P\{\boldsymbol{T} \in V_{\gamma,2}\} + P\{\boldsymbol{T} \in V_{\gamma,3}\} + P\{\boldsymbol{T} \in V_{\gamma,4}\}$ as a function of c_γ. From the definitions of $V_{\gamma,1}$, $V_{\gamma,2}$, $V_{\gamma,3}$ and $V_{\gamma,4}$, it is clear that

$$\begin{aligned} & P\{\boldsymbol{T} \in V_{\gamma,1}\} = P\{\boldsymbol{T} \in V_{\gamma,4}\} \\ = & \int_0^\gamma g \sin^{k-1}\theta \ .\ P\{R_{\boldsymbol{T}} \cos(\phi - \gamma) \leq c_\gamma\}\, d\theta \\ = & \int_0^\gamma g \sin^{k-1}\theta \ .\ F_{(k+1),\nu}\left(\frac{c_\gamma^2}{(k+1)\cos^2(\phi - \gamma)}\right) d\theta, \end{aligned}$$

and

$$\begin{aligned} & P\{\boldsymbol{T} \in V_{\gamma,2}\} = P\{\boldsymbol{T} \in V_{\gamma,3}\} \\ = & \int_\gamma^{\frac{\pi}{2}} g \sin^{k-1}\theta \ .\ P\{R_{\boldsymbol{T}} \cos(\theta - \phi) \leq c_\gamma\}\, d\theta \\ = & \int_\gamma^{\frac{\pi}{2}} g \sin^{k-1}\theta \ .\ F_{(k+1),\nu}\left(\frac{c_\gamma^2}{(k+1)\cos^2(\theta - \phi)}\right) d\theta. \end{aligned}$$

Therefore, $P\{\boldsymbol{T} \in V_\gamma\}$ is given by

$$\int_0^\gamma 2g \sin^{k-1}\theta \ . \ F_{(k+1),\nu}\left(\frac{c_\gamma^2}{(k+1)\cos^2(\phi-\gamma)}\right) d\theta$$
$$+ \quad \int_\gamma^{\frac{\pi}{2}} 2g \sin^{k-1}\theta \ . \ F_{(k+1),\nu}\left(\frac{c_\gamma^2}{(k+1)\cos^2(\theta-\phi)}\right) d\theta. \tag{4.12}$$

Expression (4.12) gives the confidence level of the inner-hyperbolic band for a given c_γ which can be used to calculate the critical constant c_γ for a given α. It is noteworthy that when $\gamma = \phi$, it matches the expression for the confidence level of the hyperbolic band given in (4.5), whereas when $\gamma = 0$, it matches the expression for the confidence level of the constant width band given in (4.7).

4.6.2 Volume of confidence set

To compute the volume of the confidence set C_γ corresponding to the inner-hyperbolic band, we use the partitioning $v(V_\gamma) = v(V_{\gamma,1}) + v(V_{\gamma,2}) + v(V_{\gamma,3}) + v(V_{\gamma,4})$.

Hence, the volume of the regions $V_{\gamma,1}$ and $V_{\gamma,4}$ are given by

$$\begin{aligned}
& v(V_{\gamma,1}) \ = \ v(V_{\gamma,4}) \\
= \quad & \iint_{\substack{R_{\boldsymbol{T}}\cos(\phi-\gamma) \le c_\gamma \\ 0<\theta_1\le\gamma}} \int_{\theta_2=0}^{\pi} \cdots \int_{\theta_{k-1}=0}^{\pi} \int_{\theta_k=0}^{2\pi} |J| dR d\theta_1 d\theta_2 \ldots d\theta_k \\
= \quad & \left\{ v\left(B_{k+1}(c_\gamma)\right) / \int_{R=0}^{c_\gamma}\int_{\theta_1=0}^{\pi} R^k \sin^{k-1}\theta_1 \, dR d\theta_1 \right\} \\
\times \quad & \iint_{\substack{R_{\boldsymbol{T}}\cos(\phi-\gamma) \le c_\gamma \\ 0<\theta_1\le\gamma}} R^k \sin^{k-1}\theta_1 \, dR d\theta_1 \\
= \quad & g \int_{\theta_1=0}^{\gamma} \frac{\sin^{k-1}\theta_1}{\cos^{k+1}(\phi-\gamma)} \, d\theta_1 \ . \ v\left(B_{k+1}(c_\gamma)\right).
\end{aligned}$$

Similarly, the volume of the regions $V_{\gamma,2}$ and $V_{\gamma,3}$ are given by

$$\begin{aligned}
& v(V_{\gamma,2}) \ = \ v(V_{\gamma,3}) \\
= \quad & \iint_{\substack{R_{\boldsymbol{T}}\cos(\theta_1-\phi) \le c_\gamma \\ \gamma<\theta_1\le\frac{\pi}{2}}} \int_{\theta_2=0}^{\pi} \cdots \int_{\theta_{k-1}=0}^{\pi} \int_{\theta_k=0}^{2\pi} |J| dR d\theta_1 d\theta_2 \ldots d\theta_k \\
= \quad & g \int_{\theta_1=\gamma}^{\frac{\pi}{2}} \frac{\sin^{k-1}\theta_1}{\cos^{k+1}(\theta_1-\phi)} \, d\theta_1 \ . \ v\left(B_{k+1}(c_\gamma)\right).
\end{aligned}$$

Therefore, the volume of V_γ is given by

$$v(V_\gamma) = 2g \left[\int_{\theta_1=0}^{\gamma} \frac{\sin^{k-1}\theta_1}{\cos^{k+1}(\phi-\gamma)} \, d\theta_1 \ + \ \int_{\theta_1=\gamma}^{\frac{\pi}{2}} \frac{\sin^{k-1}\theta_1}{\cos^{k+1}(\theta_1-\phi)} \, d\theta_1 \right] V\left(B_{k+1}(c_\gamma)\right).$$

Note that

$$\begin{aligned}
C_\gamma \ &= \ \left\{ \boldsymbol{b} : \ (\boldsymbol{z}, Z)^{-1}(X^T X)(\hat{\boldsymbol{b}} - \boldsymbol{b})/\hat{\sigma} \in V_\gamma \right\} \\
&= \ \left\{ \boldsymbol{b} : \ \boldsymbol{b} \in \hat{\boldsymbol{b}} + \hat{\sigma}(X^T X)^{-1}(\boldsymbol{z}, Z) V_\gamma \right\}
\end{aligned}$$

and so

$$v(C_\gamma) = |\hat{\sigma}(X^TX)^{-1}(\boldsymbol{z}, Z)|v(V_\gamma) = \hat{\sigma}^{k+1}|(X^TX)^{-\frac{1}{2}}|v(V_\gamma).$$

Hence, the volume of C_γ is given by

$$2\hat{\sigma}^{k+1}|(X^TX)^{-\frac{1}{2}}|g\frac{\pi^{\frac{k+1}{2}}c_\gamma^{k+1}}{\Gamma\left[\frac{k+1}{2}+1\right]}\left[\int_0^\gamma \frac{\sin^{k-1}\theta_1}{\cos^{k+1}(\phi-\gamma)}\,d\theta_1 + \int_\gamma^{\frac{\pi}{2}} \frac{\sin^{k-1}\theta_1}{\cos^{k+1}(\theta_1-\phi)}\,d\theta_1\right]. \tag{4.13}$$

When $\gamma = \phi$, expression (4.13) gives the volume of confidence set for the two-sided hyperbolic band in (4.6) and when $\gamma = 0$, it gives the volume of confidence set for the two-sided constant-width band in (4.8).

4.7 Searching for the best inner-hyperbolic band over χ_E

For given ϕ (or r), k, ν and α, we can numerically search over the family of exact $1-\alpha$ level inner-hyperbolic bands for the optimal inner-hyperbolic band that minimizes $v(C_\gamma)$ in a similar way as in Section 3.1.3. For each $\gamma \in [0, \phi]$, we determine the critical constant c_γ of the inner-hyperbolic band using (4.12) and then calculate the volume of its confidence set by using (4.13). We search over $\gamma \in [0, \phi]$ to find the $\gamma^* \in [0, \phi]$ that gives the smallest volume of confidence set. This method is illustrated below using the Snee acetylene dataset given in Table 4.1. The first two predictor variables, namely the reactor temperature x_1 and the ratio of H_2 to n-Heptane x_2, are used to illustrate the case when $k = 2$. As in Section 4.5, we assume that we wish to bound the regression function $\boldsymbol{x}^T\boldsymbol{b}$ using a $1-\alpha = 0.90$ level simultaneous confidence band over the ellipsoidal region χ_E with $r = 1.9$, so that the region of interest χ_E is given by

$$\chi_E = \left\{\boldsymbol{x}_{(1)} :\ \boldsymbol{x}^T(X^TX)^{-1}\boldsymbol{x} \leq 0.288125\right\}$$

as depicted in Figure 4.10 by the bigger ellipse in the (x_1, x_2)-plane with $\phi = 1.0863$. For each value of $\gamma \in [0, \phi]$, the corresponding critical constant c_γ of the inner-hyperbolic band is computed using expression (4.12). Then, the corresponding volume of confidence set is calculated from expression (4.13). The volume of the confidence set against γ is plotted in Figure 4.9 from which the $\gamma \in [0, \phi]$ that gives the MVCS, i.e. the best inner-hyperbolic band, can be identified. Specifically, the volume of the confidence set corresponding to the best inner-hyperbolic band is 0.6403 $units^3$, whereas those corresponding to the hyperbolic and constant width bands are 0.6419 and 0.7182 $units^3$ respectively. Furthermore, the optimal $\gamma^* \in [0, \phi]$ is given by 0.8332 rad with the corresponding $r^* = 1.1005$ from (4.9) and the critical value $c_\gamma = 2.6874$. The best inner-hyperbolic band is shown in Figure 4.10. The band has a hyperbolic shape within the region $\chi_{\gamma,E}$ depicted in Figure 4.10 by the smaller ellipse in the (x_1, x_2)-plane, whereas the band spans linearly in the region χ_{oE} that is inside χ_E but outside $\chi_{\gamma,E}$. Unfortunately, these features cannot be easily distinguished in Figure 4.10. Therefore, cross-sectional plots, Figure 4.11 and Figure 4.12, are used to show the cross-section of the band along the x_1-direction at $x_2 = x_{\cdot 2}$ and along the

Figure 4.9: Volume of confidence set C_γ against γ

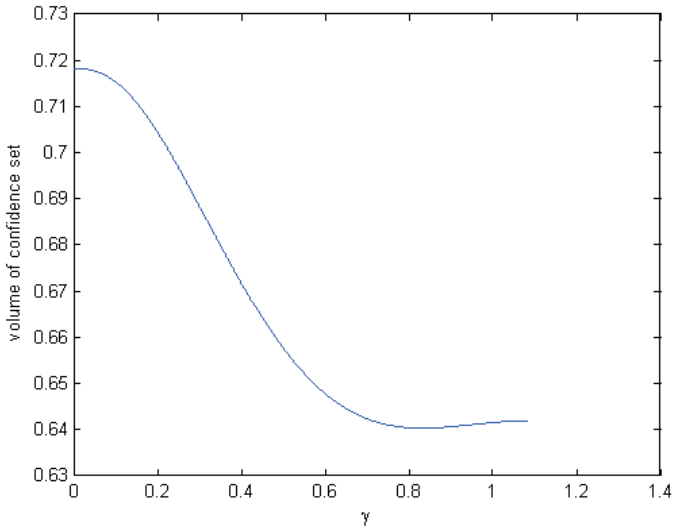


Figure 4.10: The best 0.95 level inner-hyperbolic band over χ_E, Snee acetylene dataset

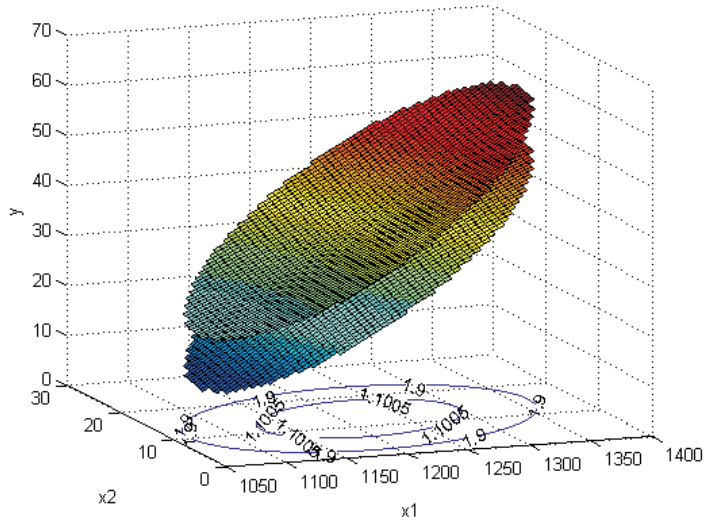


Figure 4.11: Cross-section of the best inner-hyperbolic band along the x_1-direction at $x_2 = x_{\cdot 2}$

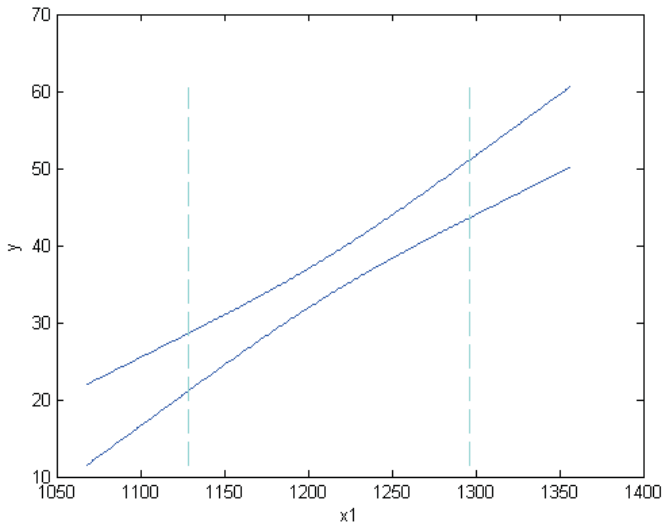


Figure 4.12: Cross-section of the best inner-hyperbolic band along the x_2-direction at $x_1 = x_{\cdot 1}$

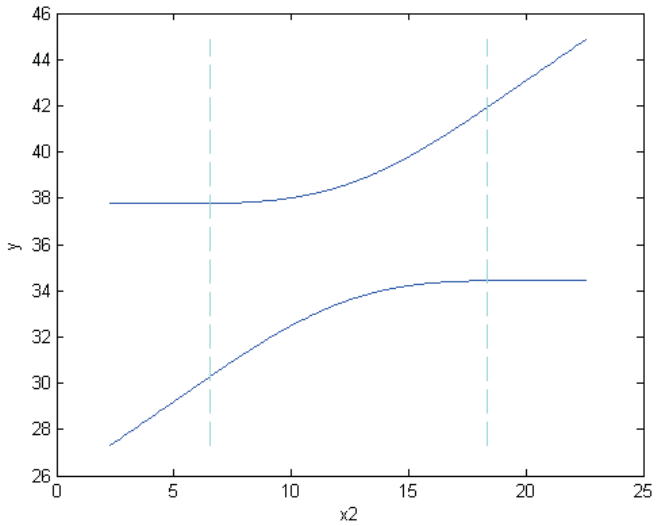
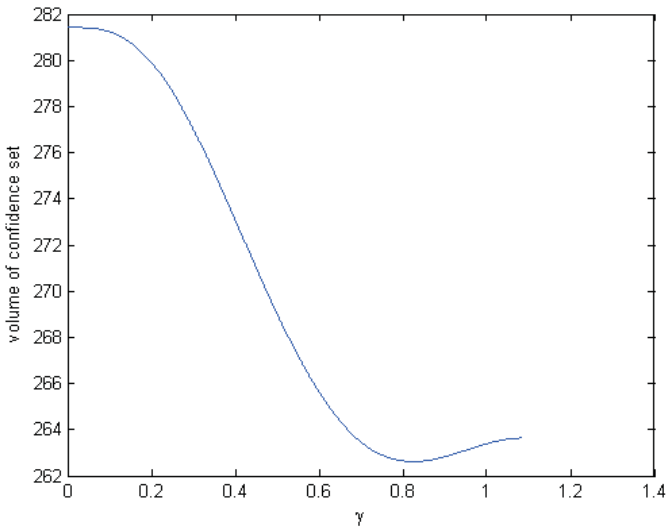


Figure 4.13: Volume of confidence set C_γ against γ when $k = 3$



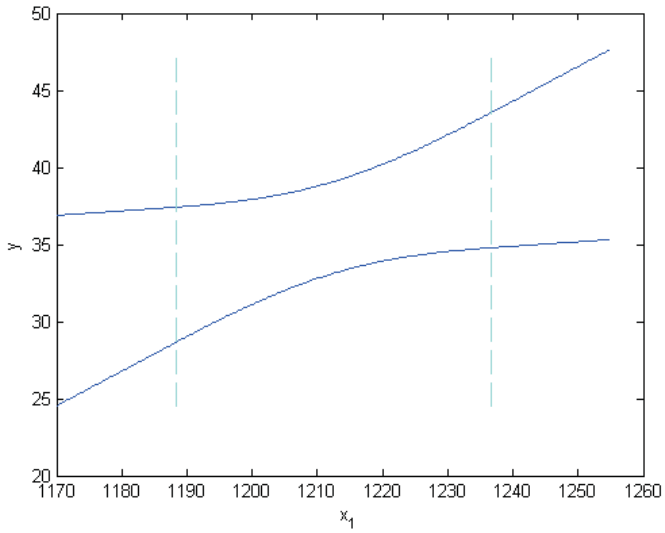
x_2-direction at $x_1 = x_{.1}$ respectively, together with the respective cross-sections of the "inner-range" formed by $\chi_{\gamma,E}$. Note that if S^{-1} is the block matrix given by

$$S^{-1} = \begin{pmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{pmatrix}$$

then the inner-range along the x_1-direction at $x_2 = x_{.2}$ is calculated to be $(x_{.1} \pm \sqrt{r_\gamma^2 / S_{11}}) = (1128.7681, 1296.2319)$. Similarly, the inner-range along the x_2-direction at $x_1 = x_{.1}$ is calculated to be $(x_{.2} \pm \sqrt{r_\gamma^2 / S_{22}}) = (6.5634, 18.3242)$.

The third predictor variable of the Snee (1977) acetylene dataset, the contact time x_3, is included to the model to illustrate the numerical search when $k = 3$. The method used to bound the regression function $\boldsymbol{x}^T\boldsymbol{b}$ using a $1 - \alpha = 0.90$ level simultaneous confidence band over the ellipsoidal region χ_E with $r = 1.9$ is the same as for $k = 2$. In this case, the volume of confidence set varies with γ as shown in Figure 4.13. The volume of the confidence set corresponding to the best inner-hyperbolic band is 262.6202 $units^4$, whereas those corresponding to the hyperbolic and constant width bands are 263.6401 and 281.4209 $units^4$ respectively. The optimal $\gamma^* \in [0, \phi]$ is given by 0.8245 rad with the corresponding $r^* = 1.10050$ from (4.9) and the critical value $c_\gamma = 3.0500$. The best inner-hyperbolic band cannot be pictured for $k = 3$, but cross-sections of the band along predictor variables can be plotted. For instance, a cross-section of the band along the x_1-direction at $x_2 = x_{.2}$ and $x_3 = x_{.3}$ is depicted in Figure 4.14, from which it can be observed that the band bounds the regression function $\boldsymbol{x}^T\boldsymbol{b}$ in $[33.1325, 39.0800]$ when $x_1 = x_{.1}$, $x_2 = x_{.2}$ and $x_3 = x_{.3}$.

Figure 4.14: Cross-section of the best inner-hyperbolic band along the x_1-direction at $x_2 = x_{.2}$ and $x_3 = x_{.3}$



4.8 Concluding remarks on the inner-hyperbolic band over χ_E

A new family of simultaneous confidence bands over an ellipsoidal covariate region has been introduced and it is shown how the best confidence band in the family can be identified numerically. It is noteworthy that the expressions for the confidence level and volume of confidence set provided in this chapter are valid for any number k of predictor variables. Given the confidence level and a design matrix X, the methods described in this chapter can be used to calculate the critical constant for the best inner-hyperbolic band and the corresponding volume of confidence set. For $k = 1$ and $k = 2$, detailed images of the band can be plotted whereas for $k > 2$, cross-sectional views of the band along the means of certain predictor variables can be generated.

The best inner-hyperbolic band can be considerably more efficient than the constant width band, as shown in the example in Section 4.7. However, the gain in efficiency in using the best inner-hyperbolic band over the hyperbolic band may be small (at least for the examples in Section 4.7). Therefore, the hyperbolic band may be recommended if one wants to avoid the numerical search to find the best inner-hyperbolic band.

In the next chapter, we consider the construction and comparison of exact confidence bands in multiple linear regression over a rectangular covariate region.

Chapter 5

Exact simultaneous confidence bands in multiple linear regression with predictor variables constrained in a rectangular region

5.1 The rectangular region

In common practice, each predictor variable in a multiple linear regression model is usually bounded and the covariate region over which a simultaneous confidence band is required is rectangular. Several authors (see e.g., Casella and Strawderman, 1980 and Naiman, 1987) have agreed that the rectangular region given by

$$\chi_R = \left\{ \boldsymbol{x}_{(1)} : \ a_i \leq x_i \leq b_i, \ i = 1, \ldots, k \right\}, \tag{5.1}$$

where $-\infty \leq a_i < b_i \leq \infty$ are given constants, is one of the most useful covariate region. Construction of conservative two-sided hyperbolic confidence bands over χ_R when k is small has been considered by Knafl, Sacks and Ylvisaker (1985), Naiman (1987, 1990) and Sun and Loader (1994) among others. More recently, a simulation-based method has been used to compute critical constants for hyperbolic bands (Liu *et al.*, 2005a) and constant width bands (Liu *et al.*, 2005b) for any given $k \geq 1$ over χ_R. They expressed a critical constant as the $1-\alpha$ population percentile of a distribution and used the $1-\alpha$ sample percentile of an *i.i.d.* sample from this distribution as an approximation to the critical constant. Apart from simulation methods, there has been no published methods for the construction of exact simultaneous confidence bands over χ_R.

In this chapter, exact $1-\alpha$ level simultaneous hyperbolic and constant width bands over χ_R are constructed by expressing the confidence level of the bands as k-dimensional integrals. The hyperbolic and constant width bands are then compared using the average width (see e.g., Naiman, 1984) and minimum volume confidence set (see e.g., Liu and Hayter, 2007 and Liu *et al.*, 2009) optimality criteria. As in previous chapters, the key is

a transformation from cartesian to polar coordinates but a further numerical quadrature is required to implement the method.

5.2 Preliminaries

Recall from previous chapters that U is the unique square root matrix of $(X^TX)^{-1}$ so that $(X^TX)^{-1} = U^2$. It immediately follows that $\boldsymbol{N} = U^{-1}(\hat{\boldsymbol{b}} - \boldsymbol{b})/\sigma \sim N_{k+1}(\boldsymbol{0}, \boldsymbol{I})$ and $\boldsymbol{T} = \boldsymbol{N}/(\frac{\hat{\sigma}}{\sigma}) = U^{-1}(\hat{\boldsymbol{b}} - \boldsymbol{b})/\hat{\sigma}$ is a standard multivariate t random vector of $k+1$ dimensions with $\nu = n - k - 1$ degrees of freedom (see e.g., Tong, 1990).

The transformation from cartesian to polar (hyperspherical) coordinates is similar to the transformation in Section 4.2 where the polar coordinates of the $(k+1)$-dimensional vector $\boldsymbol{T} = (t_1, \ldots, t_{k+1})^T$, $(R_{\boldsymbol{T}}, \theta_{\boldsymbol{T}1}, \ldots, \theta_{\boldsymbol{T}k})^T$, are defined by

$$\begin{aligned}
t_1 &= R_{\boldsymbol{T}} \cos\theta_{\boldsymbol{T}1} \\
t_2 &= R_{\boldsymbol{T}} \sin\theta_{\boldsymbol{T}1} \cos\theta_{\boldsymbol{T}2} \\
t_3 &= R_{\boldsymbol{T}} \sin\theta_{\boldsymbol{T}1} \sin\theta_{\boldsymbol{T}2} \cos\theta_{\boldsymbol{T}3} \\
&\vdots \quad\quad \vdots \\
t_k &= R_{\boldsymbol{T}} \sin\theta_{\boldsymbol{T}1} \sin\theta_{\boldsymbol{T}2} \ldots \sin\theta_{\boldsymbol{T}k-1} \cos\theta_{\boldsymbol{T}k} \\
t_{k+1} &= R_{\boldsymbol{T}} \sin\theta_{\boldsymbol{T}1} \sin\theta_{\boldsymbol{T}2} \ldots \sin\theta_{\boldsymbol{T}k-1} \sin\theta_{\boldsymbol{T}k}
\end{aligned}$$

where

$$\begin{aligned}
0 &\leq \theta_{\boldsymbol{T}1} \leq \pi \\
0 &\leq \theta_{\boldsymbol{T}2} \leq \pi \\
&\vdots \quad\quad \vdots \\
0 &\leq \theta_{\boldsymbol{T}k-1} \leq \pi \\
0 &\leq \theta_{\boldsymbol{T}k} \leq 2\pi \\
R_{\boldsymbol{T}} &\geq 0.
\end{aligned}$$

The joint density function of $(R_{\boldsymbol{T}}, \theta_{\boldsymbol{t}1}, \ldots, \theta_{\boldsymbol{t}k})^T$ can be found using the Jacobian of the transformation

$$|J| = R_{\boldsymbol{T}}^k \sin^{k-1}\theta_{\boldsymbol{T}1} \sin^{k-2}\theta_{\boldsymbol{T}2} \ldots \sin\theta_{\boldsymbol{T}k-1}.$$

However, in this case, the marginal density of $\theta_{\boldsymbol{T}j}$ $(1 \leq j \leq k-1)$ is given by

$$f_j(\theta) = g_j \sin^{k-1}\theta \quad , \quad 0 \leq \theta \leq \pi \tag{5.2}$$

where $g_j = 1/(\int_0^\pi \sin^{k-j}\theta \, d\theta)$ is the normalizing constant, the marginal density of $\theta_{\boldsymbol{T}k}$ is uniform on the interval $[0, 2\pi]$, and the marginal distribution of $R_{\boldsymbol{T}}$ is given by

$$R_{\boldsymbol{T}} \sim \sqrt{(k+1)F_{(k+1),\nu}}$$

where $F_{(k+1),\nu}$ denotes an F random variable that has $(k+1)$ and ν degrees of freedom.

5.3 Two-sided hyperbolic band over χ_R

5.3.1 Confidence level

The hyperbolic simultaneous confidence band over the region χ_R has the form

$$\boldsymbol{x}^T\boldsymbol{b} \in \boldsymbol{x}^T\hat{\boldsymbol{b}} \pm c_h\hat{\sigma}\sqrt{\boldsymbol{x}^T(X^TX)^{-1}\boldsymbol{x}} \quad \forall \boldsymbol{x}_{(1)} = (x_1, \ldots, x_k)^T \in \chi_R \tag{5.3}$$

where the critical constant c_h is chosen so that the confidence level of the band is $1-\alpha$. The confidence level of the band can be expressed as

$$\begin{aligned}
& P\left\{\sup_{x_i\in[a_i,b_i],i=1,\ldots,k} \frac{|\boldsymbol{x}^T(\hat{\boldsymbol{b}}-\boldsymbol{b})|}{\hat{\sigma}\sqrt{\boldsymbol{x}^T(X^TX)^{-1}\boldsymbol{x}}} < c_h\right\} \\
= \; & P\left\{\sup_{x_i\in[a_i,b_i],i=1,\ldots,k} \frac{|(U\boldsymbol{x})^T\boldsymbol{T}|}{\sqrt{(U\boldsymbol{x})^T(U\boldsymbol{x})}} < c_h\right\} \\
= \; & P\left\{\sup_{x_i\in[a_i,b_i],i=1,\ldots,k} \frac{|(U\boldsymbol{x})^T\boldsymbol{T}|}{\|U\boldsymbol{x}\|} < c_h\right\} \\
= \; & P\left\{\|\boldsymbol{T}\| < c_h\left(\sup_{x_i\in[a_i,b_i],i=1,\ldots,k} \frac{|(U\boldsymbol{x})^T\boldsymbol{T}|}{\|U\boldsymbol{x}\|\|\boldsymbol{T}\|}\right)^{-1}\right\} \\
= \; & P\left\{R_{\boldsymbol{T}} < c_h/Q_h\right\}
\end{aligned} \tag{5.4}$$

since $R_{\boldsymbol{T}} = \|\boldsymbol{T}\|$, where

$$Q_h = Q_h(\theta_{\boldsymbol{T}1}, \ldots, \theta_{\boldsymbol{T}k}) = \sup_{x_i\in[a_i,b_i],i=1,\ldots,k} \frac{|(U\boldsymbol{x})^T\boldsymbol{T}|}{\|U\boldsymbol{x}\|\|\boldsymbol{T}\|}.$$

The function $Q_h = Q_h(\theta_{\boldsymbol{T}1}, \ldots, \theta_{\boldsymbol{T}k})$ depends on $(\theta_{\boldsymbol{T}1}, \ldots, \theta_{\boldsymbol{T}k})^T$ only and can be quickly and accurately computed for a given $(\theta_{\boldsymbol{T}1}, \ldots, \theta_{\boldsymbol{T}k})^T$ by using a simple quadratic programming method given in Liu *et al.* (2005a) (see Appendix A). The expression (5.4) can further be expressed as

$$\begin{aligned}
& \int_{\theta_1=0}^{\pi}\cdots\int_{\theta_{k-1}=0}^{\pi}\int_{\theta_k=0}^{2\pi} P\left\{R_{\boldsymbol{T}} < c_h/Q_h | \theta_{\boldsymbol{T}1}=\theta_1, \ldots, \theta_{\boldsymbol{T}k}=\theta_k\right\} \\
& \qquad \times f\left\{\theta_{\boldsymbol{T}1}=\theta_1, \ldots, \theta_{\boldsymbol{T}k}=\theta_k\right\} d\theta_1\ldots d\theta_k \\
= \; & \int_{\theta_1=0}^{\pi}\cdots\int_{\theta_{k-1}=0}^{\pi}\int_{\theta_k=0}^{2\pi} \frac{1}{2\pi} f_1(\theta_1)\ldots f_{k-1}(\theta_{k-1}) P\left\{R_{\boldsymbol{T}} < c_h/Q_h\right\}\; d\theta_1\ldots d\theta_k \\
= \; & \int_{\theta_1=0}^{\pi}\cdots\int_{\theta_{k-1}=0}^{\pi}\int_{\theta_k=0}^{2\pi} \frac{1}{2\pi} f_1(\theta_1)\ldots f_{k-1}(\theta_{k-1}) \\
\times \; & F_{k+1,\nu}\left(c_h^2/(k+1)Q_h^2\right)\; d\theta_1\ldots d\theta_k.
\end{aligned} \tag{5.5}$$

Expression (5.5) gives the confidence level of the hyperbolic band and it involves only a k-dimensional numerical quadrature since both Q_h and $F_{k+1,\nu}(\cdot)$ can be computed quickly and accurately. For an indication of typical values and time taken under different types of numerical quadrature methods, see Appendix B. In particular, for $k=2$, the confidence level of the band is given by

$$\int_{\theta_1=0}^{\pi}\int_{\theta_2=0}^{2\pi} \frac{1}{4\pi}\sin\theta_1 F_{3,\nu}\left(\frac{c_h^2}{3(Q_h(\theta_1,\theta_2))^2}\right)\; d\theta_1 d\theta_2.$$

For $k = 3$, the confidence level is given by

$$\int_{\theta_1=0}^{\pi}\int_{\theta_2=0}^{\pi}\int_{\theta_3=0}^{2\pi} \frac{1}{2\pi^2}\sin^2\theta_1 \sin\theta_2 F_{4,\nu}\left(\frac{c_h^2}{4(Q_h(\theta_1,\theta_2,\theta_3))^2}\right)\, d\theta_1 d\theta_2 d\theta_3.$$

The confidence level of the band for $k = 2$ and $k = 3$ can be readily and accurately calculated using numerical quadrature (*e.g.* the in-built functions *dblquad* and *triplequad* in Matlab). For $k > 3$, until evaluation of higher dimensional quadrature is made possible by computer software, the simulation method used in Liu *et al.* (2005a) is recommended.

5.3.2 Average width

The width of the hyperbolic band in (5.3) is equal to $2c_h\hat{\sigma}\sqrt{\boldsymbol{x}^T(X^TX)^{-1}\boldsymbol{x}}$ at $\boldsymbol{x}_{(1)} \in \chi_R$. Hence, the average width of the band is given by

$$AW_h = \int_{x_1=a_1}^{b_1} \ldots \int_{x_k=a_k}^{b_k} \frac{2c_h\hat{\sigma}\sqrt{\boldsymbol{x}^T(X^TX)^{-1}\boldsymbol{x}}}{\prod_{j=1}^{k}(b_j - a_j)}\, dx_1 \ldots dx_k. \tag{5.6}$$

Expression (5.6) can be easily calculated for $k \leq 3$ by numerical quadrature, using for example the in-built functions *dblquad* and *triplequad* in Matlab. For $k > 3$, a simulation method can be used. Note that (5.6) can also be expressed as

$$\begin{aligned} & 2c_h\hat{\sigma}\int_{\boldsymbol{x}_{(1)}\in\chi_R} \sqrt{\boldsymbol{x}^T(X^TX)^{-1}\boldsymbol{x}}\, d\boldsymbol{x}_{(1)} \;/\; \int_{\boldsymbol{x}_{(1)}\in\chi_R} 1\, d\boldsymbol{x}_{(1)} \\ = \;& 2c_h\hat{\sigma}E(\sqrt{\boldsymbol{x}^T(X^TX)^{-1}\boldsymbol{x}}) \end{aligned}$$

where $E(\sqrt{\boldsymbol{x}^T(X^TX)^{-1}\boldsymbol{x}})$ is the expectation of $\sqrt{\boldsymbol{x}^T(X^TX)^{-1}\boldsymbol{x}}$ taken with respect to $\boldsymbol{x}_{(1)} = (x_1, \ldots, x_k)^T$. Each $x_i \sim U[a_i, b_i]$, $i = 1, \ldots, k$ and $x_1, \ldots, x_k$ are independent. Therefore, $E(\sqrt{\boldsymbol{x}^T(X^TX)^{-1}\boldsymbol{x}})$ can be approximated by simulation as follows.

- **Step 1:** independent $x_1^s, \ldots x_k^s$ are simulated each with $U[a_i, b_i]$ for $i = 1, \ldots, k$.
- **Step 2:** the value of $E^s = \sqrt{(1, x_1^s, \ldots x_k^s)(X^TX)^{-1}(1, x_1^s, \ldots x_k^s)^T}$ can be computed.
- **Step 3:** Steps 1 and 2 can be repeated K times to get $E_1^s, \ldots, E_K^s$ and $\bar{E} = \frac{1}{K}\sum_{j=1}^{K} E_j^s$ can be calculated to approximate $E(\sqrt{\boldsymbol{x}^T(X^TX)^{-1}\boldsymbol{x}})$.

The accuracy of this approximation to the average width AW_h can be gauged by the standard error given by

$$s.e(AW_h) = 2c_h\hat{\sigma}\sqrt{\sum_{j=1}^{K}(E_j - \bar{E})^2 \;/\; (K-1)K}.$$

5.3.3 Volume of confidence set

Let

$$V_h = \left\{ \boldsymbol{T} : \sup_{x_i\in[a_i,b_i], i=1,\ldots,k} \frac{|(U\boldsymbol{x})^T\boldsymbol{T}|}{\|U\boldsymbol{x}\|} < c_h \right\}.$$

Then, the confidence set for $\boldsymbol{b}$ corresponding to the hyperbolic band in (5.3) is given by

$$\begin{aligned} C_h &= \left\{ \boldsymbol{b} : \sup_{x_i \in [a_i, b_i], i=1,\ldots,k} \frac{|\boldsymbol{x}^T(\hat{\boldsymbol{b}} - \boldsymbol{b})|}{\hat{\sigma}\sqrt{\boldsymbol{x}^T(X^TX)^{-1}\boldsymbol{x}}} < c_h \right\} \\ &= \left\{ \boldsymbol{b} : \; U^{-1}\frac{(\hat{\boldsymbol{b}} - \boldsymbol{b})}{\hat{\sigma}} \in V_h \right\}. \end{aligned} \tag{5.7}$$

It is clear from Section 5.3.1 that

$$P\{\boldsymbol{b} \in C_h\} = P\{\boldsymbol{T} \in V_h\} = 1 - \alpha.$$

From expression (5.7), C_h can also be expressed as

$$C_h = \{\boldsymbol{b} : \; \boldsymbol{b} \in \hat{\boldsymbol{b}} + \hat{\sigma} U V_h\}.$$

Therefore, the volume of C_h is given by

$$v(C_h) = |\hat{\sigma}U| v(V_h) = \hat{\sigma}^{k+1} |(X^TX)^{-\frac{1}{2}}| v(V_h) \tag{5.8}$$

where $v(V_h)$ denotes the volume of V_h which is given by

$$\begin{aligned} v(V_h) &= \int_{\theta_1=0}^{\pi} \cdots \int_{\theta_{k-1}=0}^{\pi} \int_{\theta_k=0}^{2\pi} \int_{R=0}^{c_h/Q_h} |J| dR d\theta_1 d\theta_2 \ldots d\theta_k \\ &= \int_{\theta_1=0}^{\pi} \cdots \int_{\theta_{k-1}=0}^{\pi} \int_{\theta_k=0}^{2\pi} \frac{c_h^{k+1}}{(k+1)Q_h^{k+1}} \sin^{k-1}\theta_1 \sin^{k-2}\theta_2 \ldots \sin\theta_{k-1} dR d\theta_1 d\theta_2 \ldots d\theta_k, \end{aligned}$$

which can be computed for $k \leq 3$ using numerical quadrature. For $k > 3$, a simulation method can be used as in Section 5.3.2. Note that $v(V_h)$ can also be expressed as

$$\begin{aligned} &\int_{\theta_1=0}^{\pi} \cdots \int_{\theta_{k-1}=0}^{\pi} \int_{\theta_k=0}^{2\pi} \left\{ \frac{c_h^{k+1}}{(k+1)Q_h^{k+1}} \prod_{j=1}^{k} \delta_j \right\} \\ \times \; &\frac{\sin^{k-1}\theta_1}{\delta_1} \frac{\sin^{k-2}\theta_2}{\delta_2} \cdots \frac{\sin\theta_{k-1}}{\delta_{k-1}} \frac{1}{\delta_k} dR d\theta_1 d\theta_2 \ldots d\theta_k \end{aligned}$$

where $\frac{\sin^{k-1}\theta_1}{\delta_1}$, $\frac{\sin^{k-2}\theta_2}{\delta_2}$, …, $\frac{\sin\theta_{k-1}}{\delta_{k-1}}$ and $\frac{1}{\delta_k}$ are each a density function. Hence,

$$\begin{aligned} v(V_h) &= E\left(\frac{c_h^{k+1}}{(k+1)Q_h^{k+1}} \prod_{j=1}^{k} \delta_j \right) \\ &= \frac{c_h^{k+1}}{k+1} \prod_{j=1}^{k} \delta_j E(1/Q_h^{k+1}) \end{aligned}$$

where $E(1/Q_h^{k+1})$ is the expectation with respect to $(\theta_{\boldsymbol{T}1}, \ldots, \theta_{\boldsymbol{T}k})^T$ and $\prod_{j=1}^{k} \delta_j$ is given by

$$\int_{\theta_1=0}^{\pi} \cdots \int_{\theta_{k-1}=0}^{\pi} \int_{\theta_k=0}^{2\pi} \sin^{k-1}\theta_1 \sin^{k-2}\theta_2 \ldots \sin\theta_{k-1} dR d\theta_1 d\theta_2 \ldots d\theta_k.$$

The expectation $E(1/Q_h^{k+1})$ can be approximated by simulation as follows.

- **Step 1:** a vector $\boldsymbol{N} \sim N_{k+1}(\boldsymbol{0}, \boldsymbol{I})$ is simulated and a vector $\boldsymbol{T} = \boldsymbol{N}/(\frac{\hat{\sigma}}{\sigma})$ is simulated.

- **Step 2:** the polar coordinates $(\theta_{\boldsymbol{T}1}, \ldots, \theta_{\boldsymbol{T}k})^T$ of $\boldsymbol{T}$ are obtained and the value of $E^s = (1/Q_h(\theta_{\boldsymbol{T}1}, \ldots, \theta_{\boldsymbol{T}k})^{k+1})$ can be computed.

- **Step 3:** Steps 1 and 2 can be repeated K times to get $E_1^s, \ldots, E_K^s$ and $\bar{E} = \frac{1}{K}\sum_{j=1}^{K} E_j^s$ can be calculated to approximate $E(1/Q_h^{k+1})$.

The accuracy of this approximation to the volume of confidence set $v(C_h)$ can be gauged by the standard error given by

$$s.e(v(C_h)) = \hat{\sigma}^{k+1}|(X^TX)^{-\frac{1}{2}}|\frac{c_h^{k+1}}{k+1}\prod_{j=1}^{k}\delta_j\sqrt{\sum_{j=1}^{K}(E_j - \bar{E})^2 \;/\; (K-1)K}.$$

5.4 Two-sided constant width band over χ_R

5.4.1 Confidence level

The constant width simultaneous confidence band over the region χ_R has the form

$$\boldsymbol{x}^T\boldsymbol{b} \in \boldsymbol{x}^T\hat{\boldsymbol{b}} \pm c_c\hat{\sigma} \quad \forall \boldsymbol{x}_{(1)} = (x_1, \ldots, x_k)^T \in \chi_R \tag{5.9}$$

where the critical constant c_c is chosen so that the confidence level of the band is $1-\alpha$. The confidence level of the band can be expressed as

$$\begin{aligned}
& P\left\{\sup_{x_i\in[a_i,b_i],i=1,\ldots,k}\frac{|\boldsymbol{x}^T(\hat{\boldsymbol{b}}-\boldsymbol{b})|}{\hat{\sigma}} < c_c\right\} \\
= \; & P\left\{\sup_{x_i\in[a_i,b_i],i=1,\ldots,k}|(U\boldsymbol{x})^T\boldsymbol{T}| < c_c\right\} \\
= \; & P\left\{\|\boldsymbol{T}\| < c_c\left(\sup_{x_i\in[a_i,b_i],i=1,\ldots,k}\frac{|(U\boldsymbol{x})^T\boldsymbol{T}|}{\|\boldsymbol{T}\|}\right)^{-1}\right\} \\
= \; & P\left\{R_{\boldsymbol{T}} < c_c/Q_c\right\}
\end{aligned} \tag{5.10}$$

where

$$Q_c = Q_c(\theta_{\boldsymbol{T}1}, \ldots, \theta_{\boldsymbol{T}k}) = \sup_{x_i\in[a_i,b_i],i=1,\ldots,k}\frac{|(U\boldsymbol{x})^T\boldsymbol{T}|}{\|\boldsymbol{T}\|}.$$

It is clear that Q_c depends on $\boldsymbol{T}$ only through $(\theta_{\boldsymbol{T}1}, \ldots, \theta_{\boldsymbol{T}k})^T$. Note that $(U\boldsymbol{x})^T\boldsymbol{T}/\|\boldsymbol{T}\|$ is a linear function of $\boldsymbol{x}_{(1)} = (x_1, \ldots, x_k)^T$ and therefore attains its maximum or minimum over $\boldsymbol{x}_{(1)} \in \chi_R$ at one of the vertices of χ_R. In particular, χ_R has 2^k vertices given by

$$L = \left\{(l_1, \ldots, l_k)^T : \text{each } l_j \text{ is either } a_j \text{ or } b_j, 1 \le j \le k\right\}.$$

The function Q_c can therefore be expressed as

$$Q_c = \sup_{\boldsymbol{x}_{(1)}\in L}\frac{|(U\boldsymbol{x})^T\boldsymbol{T}|}{\|\boldsymbol{T}\|}$$

which can be easily computed since L has only 2^k points. By using similar derivation as in Section 5.3.1, the confidence level expression in (5.10) can further be written as

$$\int_{\theta_1=0}^{\pi}\cdots\int_{\theta_{k-1}=0}^{\pi}\int_{\theta_k=0}^{2\pi}\frac{1}{2\pi}f_1(\theta_1)\ldots f_{k-1}(\theta_{k-1})F_{(k+1),\nu}\left(c_c^2/(k+1)Q_c^2\right)\, d\theta_1\ldots d\theta_k. \tag{5.11}$$

Expression (5.11) gives the confidence level of the constant width band and involves only a k-dimensional quadrature. For $k = 2$, the confidence level of the band is given by

$$\int_{\theta_1=0}^{\pi}\int_{\theta_2=0}^{2\pi} \frac{1}{4\pi}\sin\theta_1 F_{3,\nu}\left(\frac{c_h^2}{3(Q_c(\theta_1,\theta_2))^2}\right)\, d\theta_1 d\theta_2.$$

For $k = 3$, the confidence level is given by

$$\int_{\theta_1=0}^{\pi}\int_{\theta_2=0}^{\pi}\int_{\theta_3=0}^{2\pi} \frac{1}{2\pi^2}\sin^2\theta_1\sin\theta_2 F_{4,\nu}\left(\frac{c_h^2}{4(Q_c(\theta_1,\theta_2,\theta_3))^2}\right)\, d\theta_1 d\theta_2 d\theta_3.$$

The confidence level of the band for $k = 2$ and $k = 3$ can be readily and accurately calculated using numerical quadrature such as the in-built functions *dblquad* and *triplequad* in Matlab. For $k > 3$, the simulation method used in Liu *et al.* (2005b) is recommended.

5.4.2 Average width

The width of the constant width band in (5.9) is given by $2c_c\hat{\sigma}$ and so the average width of the band is simply given by

$$AW_c = 2c_c\hat{\sigma}. \tag{5.12}$$

5.4.3 Volume of confidence set

Let

$$V_c = \left\{\boldsymbol{T} : \sup_{x_i\in[a_i,b_i],i=1,\ldots,k} |(U\boldsymbol{x})^T\boldsymbol{T}| < c_c\right\}.$$

Then the confidence set for $\boldsymbol{b}$ corresponding to the constant width band in (5.9) is given by

$$\begin{aligned} C_c &= \left\{\boldsymbol{b} : \sup_{x_i\in[a_i,b_i],i=1,\ldots,k} \frac{|\boldsymbol{x}^T(\hat{\boldsymbol{b}}-\boldsymbol{b})|}{\hat{\sigma}} < c_c\right\} \\ &= \left\{\boldsymbol{b} :\ U^{-1}\frac{(\hat{\boldsymbol{b}}-\boldsymbol{b})}{\hat{\sigma}} \in V_c\right\}. \end{aligned} \tag{5.13}$$

It is clear from Section 5.4.1 that

$$P\{\boldsymbol{b}\in C_c\} = P\{\boldsymbol{T}\in V_c\} = 1-\alpha.$$

From expression (5.13), C_c can also be expressed as

$$C_c = \{\boldsymbol{b} :\ \boldsymbol{b}\in\hat{\boldsymbol{b}}+\hat{\sigma}UV_c\}.$$

Therefore the volume of C_c is given by

$$v(C_c) = |\hat{\sigma}U|v(V_c) = \hat{\sigma}^{k+1}|(X^TX)^{-\frac{1}{2}}|v(V_c) \tag{5.14}$$

where

$$v(V_c) = \int_{\theta_1=0}^{\pi}\ldots\int_{\theta_{k-1}=0}^{\pi}\int_{\theta_k=0}^{2\pi}\frac{c_c^{k+1}}{(k+1)Q_c^{k+1}}\sin^{k-1}\theta_1\sin^{k-2}\theta_2\ldots\sin\theta_{k-1}dRd\theta_1d\theta_2\ldots d\theta_k.$$

For $k = 3$, $v(V_c)$ can be quickly and accurately computed using numerical quadrature. For $k > 3$, a similar simulation method as in Section 5.3.3 can be used. Note that $v(V_c)$ can also be expressed as

$$\int_{\theta_1=0}^{\pi} \cdots \int_{\theta_{k-1}=0}^{\pi} \int_{\theta_k=0}^{2\pi} \left\{ \frac{c_h^{k+1}}{(k+1)Q_c^{k+1}} \prod_{j=1}^{k} \delta_j \right\} \times \frac{\sin^{k-1}\theta_1}{\delta_1} \frac{\sin^{k-2}\theta_2}{\delta_2} \cdots \frac{\sin\theta_{k-1}}{\delta_{k-1}} \frac{1}{\delta_k} dR d\theta_1 d\theta_2 \ldots d\theta_k$$

where $\frac{\sin^{k-1}\theta_1}{\delta_1}$, $\frac{\sin^{k-2}\theta_2}{\delta_2}$, ..., $\frac{\sin\theta_{k-1}}{\delta_{k-1}}$ and $\frac{1}{\delta_k}$ are each a density function. Hence,

$$\begin{aligned} v(V_c) &= E\left(\frac{c_c^{k+1}}{(k+1)Q_c^{k+1}} \prod_{j=1}^{k} \delta_j\right) \\ &= \frac{c_c^{k+1}}{k+1} \prod_{j=1}^{k} \delta_j E(1/Q_c^{k+1}) \end{aligned}$$

where $E(1/Q_c^{k+1})$ is the expectation with respect to $(\theta_{\boldsymbol{T}1}, \ldots, \theta_{\boldsymbol{T}k})^T$. The expectation $E(1/Q_c^{k+1})$ can be approximated by simulation as follows.

- **Step 1:** a vector $\boldsymbol{N} \sim N_{k+1}(\boldsymbol{0}, \boldsymbol{I})$ is simulated and a vector $\boldsymbol{T} = \boldsymbol{N}/(\frac{\hat{\sigma}}{\sigma})$ is simulated.
- **Step 2:** the polar coordinates $(\theta_{\boldsymbol{T}1}, \ldots, \theta_{\boldsymbol{T}k})^T$ of $\boldsymbol{T}$ are obtained and the value of $E^s = (1/Q_c(\theta_{\boldsymbol{T}1}, \ldots, \theta_{\boldsymbol{T}k})^{k+1})$ can be computed.
- **Step 3:** Steps 1 and 2 can be repeated K times to get $E_1^s, \ldots, E_K^s$ and $\bar{E} = \frac{1}{K}\sum_{j=1}^{K} E_j^s$ can be calculated to approximate $E(1/Q_c^{k+1})$.

The accuracy of this approximation to the volume of confidence set $v(C_c)$ can be gauged by the standard error given by

$$s.e(v(C_c)) = \hat{\sigma}^{k+1} |(X^T X)^{-\frac{1}{2}}| \frac{c_c^{k+1}}{k+1} \prod_{j=1}^{k} \delta_j \sqrt{\sum_{j=1}^{K} (E_j - \bar{E})^2 \ / \ (K-1)K}.$$

5.5 Numerical examples

A portion of the Snee (1977) acetylene dataset in Table 4.1 is used to illustrate the results derived above when $k = 2$. The first two predictor variables, the reactor temperature x_1 and the ratio of H_2 to n-Heptane x_2 are considered in this example and the third predictor variable, the contact time x_3, is excluded. For the resulting design matrix X, the observed range $[a_1, b_1] \times [a_2, b_2] = [1100, 1300] \times [5.3, 23]$ and a confidence level of $1 - \alpha = 0.95$, the critical constant for the hyperbolic band is calculated using expression (5.5) to be $c_h = 3.1153$ and the 0.95 confidence level hyperbolic band over $[a_1, b_1] \times [a_2, b_2] = [1100, 1300] \times [5.3, 23]$ is depicted in Figure 5.1. In addition, the regions V_h is also depicted in Figure 5.2. The average width for the hyperbolic band is found to be $AW_h = 8.9338$ from expression (5.6) and the volume of confidence set corresponding to the hyperbolic band is found to be $v(C_h) = 0.2507\ units^3$ from expression (5.8). The

Figure 5.1: The 0.95 level hyperbolic band, Snee (1977) acetylene dataset, k=2

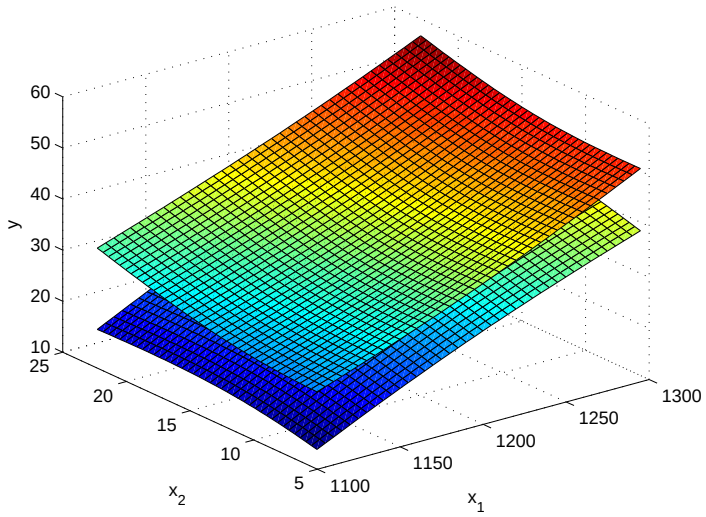


Figure 5.2: The region V_h

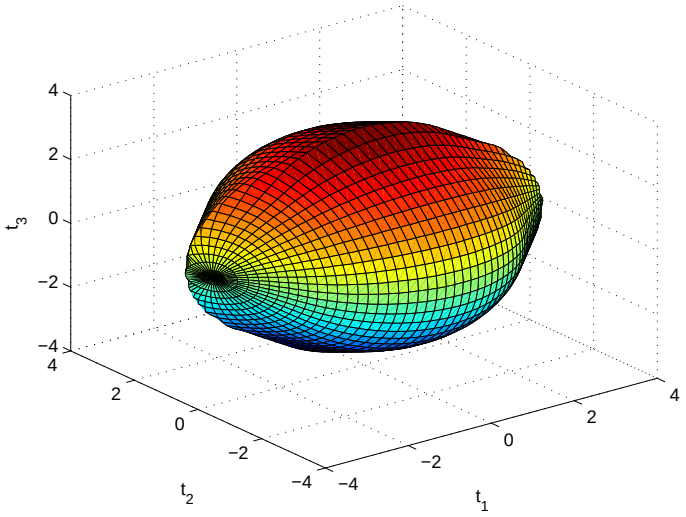
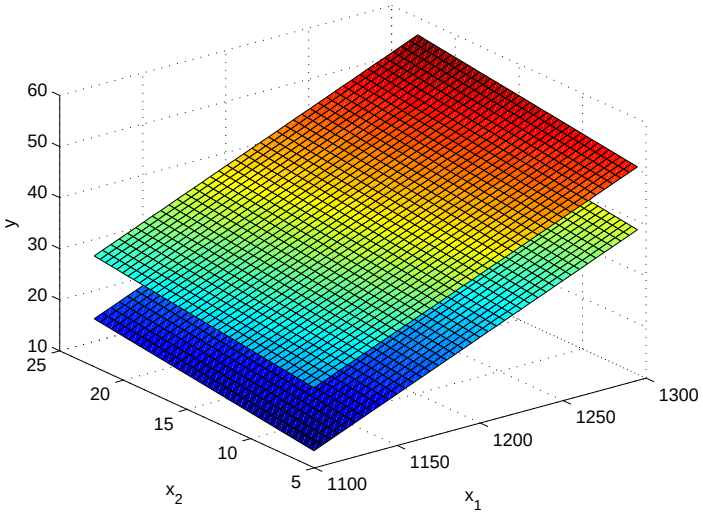


Figure 5.3: The 0.95 level constant width band, Snee (1977) acetylene dataset, k=2



simulation method has also been used to provide a means of comparison to the method of numerical quadrature. With 1×10^5 simulations, the average width for the hyperbolic band is found to be $AW_h = 8.9356$ with a standard error $s.e(AW_h) = 6.6366 \times 10^{-3}$ and the volume of confidence set corresponding to the hyperbolic band is found to be $v(C_h) = 0.2508\ units^3$ with a standard error $s.e(v(C_h)) = 1.3334 \times 10^{-4}$.

Similarly, the critical constant for the constant width band is calculated using expression (5.11) to be $c_c = 1.6984$ and the 0.95 confidence level constant width band over the same covariate region is depicted in Figure 5.3. The region V_c is also depicted in Figure 5.4. The average width for the constant width band is found to be $AW_c = 12.3099$ from expression (5.12) and the volume of confidence set corresponding to the constant width band is found to be $v(C_c) = 0.3513\ units^3$ from expression (5.14). With 1×10^5 simulations, the volume of confidence set corresponding to the constant width band is found to be $v(C_c) = 0.3527\ units^3$ with a standard error $s.e(v(C_c)) = 7.1608 \times 10^{-4}$ using the simulation method. Therefore, the hyperbolic band is better than the constant width band over the the observed range $[a_1, b_1] \times [a_2, b_2] = [1100, 1300] \times [5.3, 23]$ under both the average width criterion and the minimum volume confidence set criterion for this example. Note that the volumes of confidence sets in (5.8) and (5.14) are both of the same form and therefore the comparison between $v(V_h)$ and $v(V_c)$ can be used to compare between $v(C_h)$ and $v(C_c)$. For the reader's interest, a superposition of the regions V_h and V_c is depicted in Figure 5.5, where it is clear that the region V_h is smaller than V_c.

The third predictor variable of the Snee (1977) acetylene dataset, the contact time x_3, is included to illustrate the results when $k = 3$. For the resulting design matrix X, the observed range $[a_1, b_1] \times [a_2, b_2] \times [a_3, b_3] = [1100, 1300] \times [5.3, 23] \times [0.0115, 0.098]$ and a confidence level of $1 - \alpha = 0.95$, the critical constants for the hyperbolic and constant width bands are found to be $c_h = 3.5286$ and $c_c = 6.1614$ respectively. Their respective

Figure 5.4: The region V_c

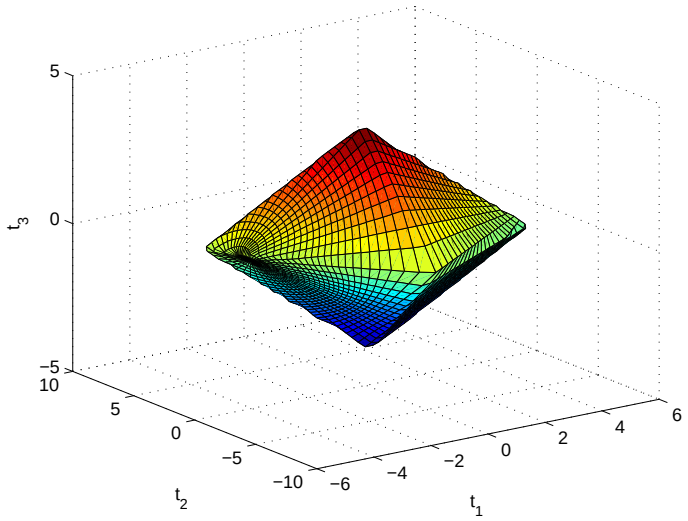
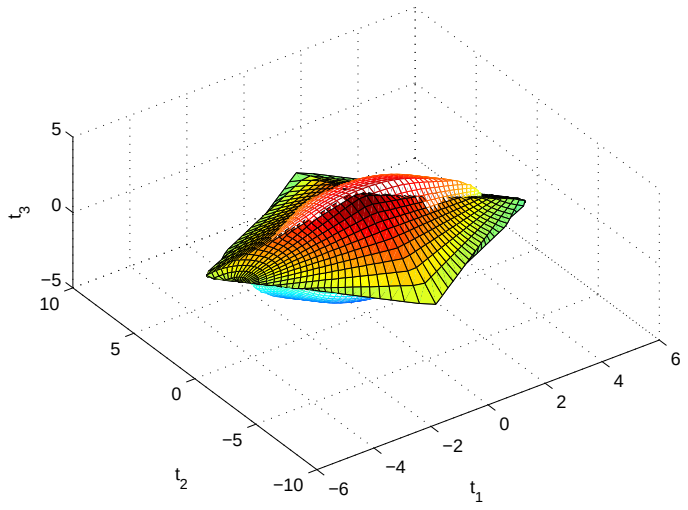


Figure 5.5: A superposition of regions V_h and V_c



average widths are calculated to be $AW_h = 25.116$ and $AW_c = 46.421$. Finally, their respective volumes of confidence sets are found to be $v(C_h) = 187.203\ units^4$ and $v(C_c) = 5055.053\ units^4$. When the simulation method is used, with 1×10^5 simulations, the average width for the hyperbolic band is found to be $AW_h = 25.0614$ with a standard error $s.e(AW_h) = 0.0419$ and the volumes of confidence sets corresponding to the hyperbolic and constant width bands are found to be $v(C_h) = 187.121\ units^4$ with a standard error $s.e(v(C_h)) = 0.1240$ and $v(C_c) = 4963.546\ units^4$ with a standard error $s.e(v(C_c)) = 38.2770$ respectively. Therefore, it is clear that the hyperbolic band is more efficient than the constant width band for this example under both the average width and minimum volume confidence set optimality criteria.

5.6 Concluding remarks on exact confidence bands over χ_R

For linear regression with $k(> 1)$ over a rectangular covariate region, only conservative or approximate methods are available in the statistical literature. A general formula for the construction of exact hyperbolic and constant width bands for k covariates restricted in intervals is provided in this chapter. The key is a transformation from cartesian to polar coordinates as in the previous chapters. To implement the method, a numerical quadrature is also necessary. For $k = 2$ and $k = 3$, the confidence levels, the required critical constants, average widths and volumes of confidence sets can be quickly and exactly computed using numerical quadrature such as the built-in functions *dblquad* and *triplequad* in Matlab. For the computation of confidence levels when $k > 3$, simulations methods are provided by Liu *et al.* (2005a) and Liu *et al.* (2005b). For the computation of average widths and volumes of confidence sets when $k > 3$, similar simulation methods have been proposed here.

The computations for the hyperbolic band are typically more time-consuming than for the constant width band due to the quadratic programming method pointed out in Section 5.3.1. However, for a rectangular covariate region χ_R, the hyperbolic band can be considerably more efficient than the constant width band under both the average width and minimum volume confidence set optimality criteria (at least for the examples in Section 5.5) and is therefore the recommended band.

Chapter 6

Conclusions and Future work

Simultaneous confidence bands in linear regression analysis are useful tools that can be applied to many aspects of real life. This thesis is a concise account of the construction of exact two-sided confidence bands in linear regression. The key method used involves a transformation from cartesian to polar coordinates and expressing the confidence level of a band as a k-dimensional integration. In simple linear regression, it has been shown analytically that a D-optimal design leads to the the best two-segment band under the MACS criterion. Attempts by numerical search to show that D-optimal designs also lead to the best three-segment and hyperbolic bands have been made. Two new families of confidence bands, called the inner-hyperbolic bands and the outer-hyperbolic bands, have been introduced and it has been shown that the best confidence band in each family can be more efficient than the best three-segment and hyperbolic bands. It is also shown numerically that the best inner-hyperbolic band is always as good as or better than the best outer-hyperbolic band. Thus, the inner-hyperbolic family of confidence bands has also been constructed in multiple linear regression over an ellipsoidal covariate region where comparisons to the hyperbolic and constant width bands have led to similar results as in simple linear regression. In the case where the predictor variables are constrained in a rectangular covariate region, a method to construct and compare between exact two-sided hyperbolic and constant width bands has been proposed for the first time.

Moreover, this thesis also points out some problems that might be of interest for future research. Although only exact two-sided confidence bands have been considered, exact one-sided bands can also be constructed using similar methods. However, the confidence sets for one-sided bands require a different interpretation as they have infinite volumes and recently, Liu *et al.* (2009) have provided a method to overcome this issue. Hence, the analytical deduction that D-optimal designs lead to the best three-segment and hyperbolic bands under under the MACS criterion is yet to be considered. Although the numerical exploration in Section 2.5 gives some insight into how the area of confidence set varies with the mean of the predictor variable $\bar{x}$ and $\sqrt{|X^TX|}$, it is not a complete answer. The relationship between $\bar{x}$ and $\sqrt{|X^TX|}$ is more complicated and thus the region over which Figure 2.21 and Figure 2.22 are plotted is not appropriate. In addition, due to its desirable properties, the outer-hyperbolic family of confidence bands can also be constructed in

multiple linear regression over an ellipsoidal covariate region, although the best band in the family is not as efficient as the best inner-hyperbolic band under the MACS criterion. When the covariate region is rectangular, although expressions for the construction of exact confidence bands have been provided, their computation is possible for up to three predictor variables so far. For $k > 3$, simulation methods are recommended until further improvement in software or future research make exact construction of confidence bands over a rectangular covariate region possible. Matlab scripts and functions have been used for all the numerical computation and illustrative plots throughout the thesis and are available upon request for the reader to explore the methods and results of the thesis. Making the computer codes available in other statistical software could also be useful future work.

The most important contribution of this thesis is the method to construct and compare exact two-sided confidence bands over a rectangular covariate region. Before this, only methods to construct conservative confidence bands or simulation methods were available in linear regression.

The hyperbolic band is a safe recommendation with only a small loss of efficiency and being generally easy to construct. However, a small increase in efficiency can have a big importance in real life situations and thus another important contribution of the thesis is the introduction of the family of inner-hyperbolic bands for simple linear regression and multiple linear regression over ellipsoidal regions and the method to construct the best band within the family.

Bibliography

1. Al-Saidy, O.M., Piegorsch, W.W., West, R.W. and Nitcheva, D.K. (2003). Confidence bands for low-dose risk estimation with quantal response data. *Biometrics, 59, 1056-1062.*

2. Atkinson, A.C., Donev, A.N. and Tobias, R.D (2007). Optimum Experimental Designs, with SAS. *Oxford University Press.*

3. Bohrer, R. (1973). A multivariate t probability integral. *Biometrika, 60, 647-654.*

4. Bowden, D.C. and Graybill, F.A. (1966). Confidence bands of uniform and proportional width for linear models. *Journal of the American Statistical Association, 61, 182-198.*

5. Casella, G. and Strawderman, W.E. (1980). Confidence bands for linear-regression with restricted predictor variables. *Journal of the American Statistical Association, 75, 862-868.*

6. Gafarian A.V. (1964). Confidence bands in straight line regression. *Journal of the American Statistical Association, 59, 182-213.*

7. Gauss C.F. (1809). Theoria Motus Corporum Coelestium in Sectionibus Conicis Solem Ambientum.

8. Graybill, F.A. and Bowden, D.C. (1967). Linear segment confidence bands for simple linear regression models. *Journal of the American Statistical Association, 62, 403-408.*

9. Halperin, M. and Gurian, J. (1968). Confidence bands in linear regression with constraints on independent variables. *Journal of the American Statistical Association, 63, 1020-1027.*

10. Khorasani, F. and Milliken, G.A. (1979). On the exactness of confidence bands about a linear model. *Journal of the American Statistical Association, 74, 446-448.*

11. Knafl, G., Sacks, J. and Ylvisaker, D. (1985). Confidence bands for regression-functions. *Journal of the American Statistical Association, 80, 683-691.*

12. Legendre A.M. (1805) Nouvelles méthodes pour la détermination des orbites des comètes.

13. Liu, W. (2010) Simultaneous inference for regression. *Forthcoming.*

14. Liu, W., Bretz, F., Hayter, A.J. and Wynn, H.P. (2009) Assessing non-superiority, non-inferiority or equivalence when comparing two regression models over a restricted covariate region. *Biometrics, to appear.*

15. Liu, W. and Hayter, A.J. (2007). Minimum area confidence set optimality for confidence bands in simple linear regression. *Journal of the American Statistical Association, 102, 181-190*

16. Liu, W., Hayter, A.J., Piegorsch, W.W. and Ah-Kine, P. (2009). Comparison of hyperbolic and constant width simultaneous confidence bands in multiple linear regression under MVCS criterion. *Journal of Multivariate Analysis, 100, 1432-1439.*

17. Liu, W., Jamshidian, M. and Zhang, Y. (2004). Multiple comparison of several regression models. *Journal of the American Statistical Association, 99, 395-403*

18. Liu, W., Jamshidian, M., Zhang, Y. and Bretz, F.(2005). Constant width simultaneous confidence bands in multiple linear regression with predictor variables constrained in intervals. *Journal of Statistical Computation and Simulation, 75(6), 425-436.*

19. Liu, W., Jamshidian, M., Zhang, Y. and J. Donnelly (2005). Simulation-based simultaneous confidence bands in multiple linear regression with predictor variables constrained in intervals. *Journal of Computational and Graphical Statistics, 14(2), 459484.*

20. Liu, W. and Lin, S. (2009). Construction of Exact Simultaneous Confidence Bands in Multiple Linear Regression with Predictor Variables Constrained in an Ellipsoidal Region. *Statistica Sinica , 19, 213-232.*

21. Liu, W., Lin, S. and Piegorsch, W.W. (2008). Construction of exact simultaneous confidence bands for a simple linear regression model. *International Statistical Review, 76, 39-57.*

22. Naiman, D.Q (1984). Average width optimality of simultaneous confidence bounds. *The Annals of Statistics, 12, 1199-1214*

23. Naiman, D.Q. (1987). Simultaneous Confidence Bounds in Multiple Regression Using Predictor Variable Constraints. *Journal of the American Statistical Association, 82, 214-219.*

24. Naiman, D.Q. (1990). Volumes of Tubular Neighborhoods of Spherical Polyhedra and Statistical Inference. *The Annals of Statistics, 18, 685-716.*

25. Piegorsch, W.W. (1987). On confidence bands and set estimators for the simple linear model. *Statistics and Probability Letters, 5(6), 409-413*

26. Piegorsch, W.W., West R.W., Pan,W. and Kodell, R. (2005). Low dose risk estimation via simultaneous confidence inferences. *Journal of the Royal Statistical Society, C, 54, 245-258*

27. Scheffé H. (1953). A method for judging all contrasts in analysis of variance. *Biometrika, 40, 87-104.*

28. Seppanen, E. and Uusipaikka, E. (1992). Confidence bands for linear-regression over restricted regions. *Scandinavian Journal of Statistics, 19, 73-81.*

29. Snee, R.D. (1977). Validation of regression models: methods and examples. *Technometrics, 19, 415-428.*

30. Spurrier, J.D., (1999). Exact confidence bounds for all contrasts of three or more regression lines *Journal of the American Statistical Association, 94, 483-488.*

31. Sun, J. and Loader, C.R. (1994.) Simultaneous confidence bands for linear regression and smoothing *The Annals of Statistics, 22, 1328-1346.*

32. Tong, Y.L. (1990). Multivariate normal distribution *New York: Springer Verlag.*

33. Uusipaikka, E. (1983). Exact confidence bands for linear-regression over intervals. *Journal of the American Statistical Association, 78, 638-644.*

34. Working, H. and Hotelling, H. (1929). Applications of the theory of error to the interpretation of trends. *Journal of the American Statistical Association, 24, 73-85.*

35. Wynn, H.P. and Bloomfield, P. (1971). Simultaneous confidence bands in regression analysis. *Journal of the Royal Statistical Society, B, 33, 202-217.*

Appendix A

Computation of $Q_h = Q_h(\theta_{\boldsymbol{T}1}, \ldots, \theta_{\boldsymbol{T}k})$

We need to compute the function

$$\begin{aligned} Q_h &= \sup_{x_i \in [a_i, b_i], i=1,\ldots,k} \frac{|(U\boldsymbol{x})^T \boldsymbol{T}|}{\|U\boldsymbol{x}\| \|\boldsymbol{T}\|} \\ &= \sup_{\boldsymbol{v} \in L(U, \chi_R)} \frac{|\boldsymbol{v}^T \boldsymbol{T}|}{\|\boldsymbol{v}\| \|\boldsymbol{T}\|} \end{aligned}$$

where $L(U, \chi_R)$ forms a cone spanned by the vectors given by $U\boldsymbol{x} = \boldsymbol{u}_0 + l_1\boldsymbol{u}_1 + \ldots + l_k\boldsymbol{u}_k$, where each l_j is either a_j or b_j, $1 \leq j \leq k$. Let $\psi(\boldsymbol{t}, U, \chi_R)$ denote the projection of $\boldsymbol{t} \in \Re^{k+1}$ to the cone $L(U, \chi_R)$. Then, it follows from Naiman (1987, Theorem 2.1) that

$$Q_h = max\left\{\|\psi(\boldsymbol{T}/\|\boldsymbol{T}\|, U, \chi_R)\|, \|\psi(-\boldsymbol{T}/\|\boldsymbol{T}\|, U, \chi_R)\|\right\}.$$

Note that $\psi(\boldsymbol{t}, U, \chi_R)$ solves the problem

$$\begin{aligned} & min_{\boldsymbol{v} \in L(U, \chi_R)} \|\boldsymbol{v} - \boldsymbol{t}\|^2 \\ = \; & min_{\boldsymbol{v} \in L(U, \chi_R)} (\boldsymbol{v}^T\boldsymbol{v} - 2\boldsymbol{t}^T\boldsymbol{v} + \boldsymbol{t}^T\boldsymbol{t}) \\ = \; & min_{\boldsymbol{v} \in L(U, \chi_R)} (\boldsymbol{v}^T\boldsymbol{v} - 2\boldsymbol{t}^T\boldsymbol{v}). \end{aligned}$$

Thus, the function $\psi(\boldsymbol{t}, U, \chi_R)$ is the vector $\boldsymbol{v} \in \Re^{k+1}$ that minimizes the function $\boldsymbol{v}^T\boldsymbol{v} - 2\boldsymbol{t}^T\boldsymbol{v}$, which is equivalent to minimizing the function

$$\frac{1}{2}\boldsymbol{v}^T\boldsymbol{v} - \boldsymbol{t}^T\boldsymbol{v},$$

subject to $\boldsymbol{v} \in L(U, \chi_R)$. From the definition of $L(U, \chi_R)$, $\boldsymbol{v} \in L(U, \chi_R)$ implies that $\boldsymbol{v} = \lambda U\boldsymbol{x}$ or equivalently

$$U^{-1}\boldsymbol{v} = \lambda\boldsymbol{x} = (\lambda, \lambda x_1, \ldots, \lambda x_k)^T \text{ for } \boldsymbol{x}_{(1)} \in \chi_R \text{ and } \lambda \geq 0.$$

Let the vector $\boldsymbol{q}_j \in \Re^{k+1}$ have the j^{th} element equal to 1 with all the remaining elements equal to 0. Then, $\boldsymbol{q}_1^T U^{-1}\boldsymbol{v} = \lambda \geq 0$ and $a_j \leq \boldsymbol{q}_{j+1}^T U^{-1}\boldsymbol{v}/\boldsymbol{q}_1^T U^{-1}\boldsymbol{v} \leq b_j$ for $j = 1, \ldots, k$ or equivalently

$$\begin{aligned} -\boldsymbol{q}_1^T U^{-1}\boldsymbol{v} &\leq 0 \\ (\boldsymbol{q}_{j+1}^T - b_j\boldsymbol{q}_1^T)U^{-1}\boldsymbol{v} &\leq 0 \text{ for } j = 1, \ldots, k \\ (a_j\boldsymbol{q}_1^T - \boldsymbol{q}_{j+1}^T)U^{-1}\boldsymbol{v} &\leq 0 \text{ for } j = 1, \ldots, k. \end{aligned}$$

These constraints can be expressed as $A\boldsymbol{v} \leq \mathbf{0}$ where the $(2k+1) \times (k+1)$ matrix A is given by

$$\begin{pmatrix} (\boldsymbol{q}_2^T - b_1\boldsymbol{q}_1^T)U^{-1} \\ (a_1\boldsymbol{q}_1^T - \boldsymbol{q}_2^T)U^{-1} \\ \vdots \\ (\boldsymbol{q}_{k+1}^T - b_k\boldsymbol{q}_1^T)U^{-1} \\ (a_k\boldsymbol{q}_1^T - \boldsymbol{q}_{k+1}^T)U^{-1} \\ -\boldsymbol{q}_1 U^{-1} \end{pmatrix}.$$

The problem of minimizing the function $\frac{1}{2}\boldsymbol{v}^T\boldsymbol{v} - \boldsymbol{t}^T\boldsymbol{v}$ under the constraints $A\boldsymbol{v} \leq \mathbf{0}$ is a standard quadratic programming problem that can be solved numerically (*e.g.* by Matlab using the in-built function *quadprog(H,f,A,b)*). Therefore, given $\theta_{\boldsymbol{T}1}, \ldots, \theta_{\boldsymbol{T}k}$, the unit vector $\boldsymbol{T}/\|\boldsymbol{T}\|$ can be calculated and the value of $\psi(\boldsymbol{T}/\|\boldsymbol{T}\|, U, \chi_R)$ can be computed using, for example, *quadprog(H,f,A,b)* in Matlab to obtain Q_h.

Appendix B

Matlab computation values and times for confidence bands over χ_R

The computation of critical constants, average widths and volume of confidence sets for a given dataset (or design matrix X) and confidence level $1-\alpha$ in the thesis have been performed on Matlab. The scripts and functions produced for the contents of this thesis are all available upon request.

The computations involved in Chapter 5 are typically time consuming. Computations for the previous chapters are very quick and accurate, taking usually less than a minute. However, the implementation of the quadratic programming problem in the numerical quadratures in Chapter 5 is more computer intensive. Tables B.1 - B.6 show the computation values and time taken for various methods of numerical quadrature and tolerance levels for $k = 2$ and $k = 3$ on a dual core PC, 3.0GHz, 3.0GHz, 1.99 GB RAM. Two levels of tolerances are used, $tol = 10^{-6}$ and $tol = 10^{-3}$. The parameter tol refers to the absolute error tolerance used in Matlab for numerical quadratures and the default value is $tol = 10^{-6}$. Larger values of tol result in fewer function evaluations and faster computation, but less accurate results. Two methods of numerical quadratures are used, *quad* and *quadl*. The function *quad* refers to a recursive adaptive Simpson quadrature used in Matlab. It is most efficient for low accuracies with non-smooth integrands. The function *quadl* refers to a recursive adaptive Lobatto quadrature used in Matlab. It is more efficient than *quad* for high accuracies with smooth integrands.

The time taken for the computation of confidence levels given the values of critical constants and α are typically one tenth of the time taken for computation of critical constants given confidence levels and α. Furthermore, simulation methods are typically more computer intensive than the k-dimensional quadrature methods used in Chapter 5. Specifically, for $k = 2$, 100000 simulations took a total time taken of 470 seconds to attain a value of $v(C_h) = 0.2508$. 100000 simulations took a total time taken of 13 seconds to attain a value of $v(C_c) = 0.3527$. For $k = 3$, 100000 simulations took a total time taken of 600 seconds to attain a value of $v(C_h) = 187.121$. 100000 simulations took a total time taken of 15 seconds to attain a value of $v(C_c) = 4963.546$.

Table B.1: Computation values and times, Snee dataset, $k = 2$, $\alpha = 0.05$, $tol = 10^{-6}$, *quad*

	Computation value	Time taken (s)
c_h	3.1153	80
c_c	1.6984	12
$v(C_h)$	0.2507	122
$v(C_c)$	0.3513	7
AW_h	8.9338	0.5
AW_c	12.3099	instant

Table B.2: Computation values and times, Snee dataset, $k = 2$, $\alpha = 0.05$, $tol = 10^{-3}$, *quad*

	Computation value	Time taken (s)
c_h	3.1166	8
c_c	1.6941	0.8
$v(C_h)$	0.2510	8
$v(C_c)$	0.3487	0.5
AW_h	8.9374	instant
AW_c	12.2786	instant

Table B.3: Computation values and times, Snee dataset, $k = 2$, $\alpha = 0.05$, $tol = 10^{-3}$, *quadl*

	Computation value	Time taken (s)
c_h	3.1154	51
c_c	1.6980	3
$v(C_h)$	0.2507	55
$v(C_c)$	0.3510	3
AW_h	8.9340	instant
AW_c	12.3067	instant

Table B.4: Computation values and times, Snee dataset, $k = 3$, $\alpha = 0.05$, $tol = 10^{-6}$, *quad*

	Computation value	Time taken (s)
c_h	3.5286	2600
c_c	6.1614	1950
$v(C_h)$	187.203	23510
$v(C_c)$	5055.053	22710
AW_h	25.116	90
AW_c	46.421	instant

Table B.5: Computation values and times, Snee dataset, $k = 3$, $\alpha = 0.05$, $tol = 10^{-3}$, *quad*

	Computation value	Time taken (s)
c_h	3.5375	125
c_c	6.1602	45
$v(C_h)$	187.207	350
$v(C_c)$	5055.064	350
AW_h	25.116	1.5
AW_c	46.421	instant

Table B.6: Computation values and times, Snee dataset, $k = 3$, $\alpha = 0.05$, $tol = 10^{-3}$, *quadl*

	Computation value	Time taken (s)
c_h	3.5295	1600
c_c	6.1573	219
$v(C_h)$	187.203	7744
$v(C_c)$	5055.046	3950
AW_h	25.116	2
AW_c	46.421	instant