國立成功大學
機械工程學系
碩　士　論　文

針對機率拘束空間內最大可靠度問題之EGO演算法修正

A Modified Efficient Global Optimization Algorithm for Maximal Reliability within a Probabilistic Constrained Space

研 究 生：黃彥智
指導教授：詹魁元 博士

中 華 民 國 九 十 八 年 七 月

針對機率拘束空間內最大可靠度問題之EGO演算法修正

A Modified Efficient Global Optimization Algorithm for Maximal Reliability within a Probabilistic Constrained Space

研 究 生：黃彥智　　Student: Yen-Chih Huang

指導教授：詹魁元 博士　　Advisor: Dr. Kuei-Yuan Chan

國立成功大學

機械工程學系

碩士論文

A Thesis

Submitted to Department of Mechanical Engineering

National Cheng Kung University

in Partial Fulfillment of the Requirements

for the Degree of

Master of Science

in

Dept. of Mech. Eng.

June 2009

Tainan, Taiwan

中華民國九十八年七月

針對機率拘束空間內最大可靠度問題之EGO演算法修正

學生：黃彥智　　指導教授：詹魁元 博士

國立成功大學機械工程學系

摘　　要

最佳化設計以有系統的方式，找尋拘束空間內目標函數的最大或最小值，這樣的最佳解雖可使設計工程師了解產品的極致性能表現，但也時常位於拘束條件邊界上。任何工程問題上的不確定因素，均很可能導致最佳值滿足拘束條件的可靠度偏低。可靠度最佳化設計便是一個在設計初期便考量不確定因素的設計方法，其嚴謹卻實用的數學模式在工程領域被廣泛的應用。在過去大多數的研究中，都是在設定一個固定的可靠度作爲拘束條件來進行設計，然而對於一個未知的問題來說，訂定一個不適當的可靠度機率值，將有可能會導致問題的可行解空間並不存在，或是高估了問題本身所能要求的可靠度。因此，本論文將使用Efficient Global Optimization(EGO) 演算法，在機率形式的拘束條件下，以最大可靠度作爲目標來進行最佳化設計。

EGO演算法使用本論文所提出的取樣指標，結合其他文獻中舊有的指標，制定一個探索式的方法來不斷地在拘束條件邊界上取樣，以改善由蒙地卡羅法所計算的拘束條件之準確性，並建立Kriging模型來對各個不連續的可行解空間，進行最大可靠度的全域最佳化設計。由本論文之結果可以得知，在針對黑盒子函數問題與非連續可行解空間的可靠度問題中，此方法將比梯度型演算法或DIRECT演算法來的有效率。文末將以兩個數學範例與一個工程範例來說明演算法之結果。

A Modified Efficient Global Optimization Algorithm for Maximal Reliability within a Probabilistic Constrained Space

Student: Yen-Chih Huang Advisor: Dr. Kuei-Yuan Chan

Department of Mechanical Engineering
National Cheng Kung University

ABSTRACT

Design optimization problems under random uncertainties are commonly formulated with constraints in probabilistic forms. This formulation is also referred to as reliability-based design optimization (RBDO) in the literature and has gained extensive attentions in recent years. Most research assumes that reliability levels are given based on past experiences or other design considerations without exploring the constrained space. As a result, inappropriate target reliability levels might be assigned, which either result in null probabilistic feasible space or performances underestimation. In this research we investigate the maximal reliability withnin a probabilistic constrained space using modified efficient global optimization (EGO) algorithm. By constructing and improving Kriging models iteratively, EGO can obtain a global optimum of a possibly disconnected feasible space at high reliability levels. An infill sampling criterion (ISC) is proposed to enforce added samples on constraint boundaries to engance the accuracy of probabilistic constraint evaluations via Monte Carlo simulation. This limit state ISC, combined with existing ISC from the literature, forms a heuristic approach that efficiently improves the Kriging models. For optimization problems with expensive functions and disconnected feasible space, such as the maximal reliability problems in RBDO, the proposed work shows high efficiency and potential in finding the optimum compared to existing gradient-based and direct search methods. Several examples are used to demonstrate the proposed methodoogy.

誌　　謝

感謝恩師 詹魁元教授對於學生的細心指導與惇惇教誨，讓學生在課業與待人處事上皆獲益良多，使得學生在這兩年研究所修業的期間，成爲學生目前人生中最重要的一段時光，對於老師的教誨，學生將永誌於心。

感謝李孟穎博士在學生的研究期間，給予學生許多觀念上的指導與修正，使得學生的研究能夠有所突破。感謝陳家豪教授對於本論文的指導與建議，使得學生的論文更臻完整。感謝Sasena博士提供演算法的程式碼，讓學生可以順利學習許多演算法的技巧與概念。

感謝已畢業的學長與學弟：智豪、岱璟、淳民，在碩一時給予我的指導與照顧。感謝『系統最佳化實驗室』的夥伴：凱勛、東信、昱達、琇雯、子頡、勝昌、佳豪、季儒、祐伸，謝謝你們的扶持與幫助，陪我度過這些快樂與煎熬的時間。感謝『創意性機構設計教研室』的同學：芳富、俊榮、于庭、峻嘉、世同、惠晴、羽薰、芷誼、冠辰、明哲，謝謝你們在這兩年來總是帶給我許多的歡樂。感謝大學時代以來的好友：建宏，繼續陪我度過這兩年的研究所生活，並給予我許多生活上的幫助。

最後僅以本論文獻給我最親愛的父母親與姊姊，感謝你們多年來的支持與照顧，使我能在求學的過程中無憂無慮的完成學業。

目　　錄

書名頁 i

中文摘要 ii

英文摘要 iii

誌謝 iv

目錄 v

表目錄 viii

圖目錄 ix

符號說明 xii

第一章、緒論 (Introduction) 1

1.1 研究動機與目的 (Motivation and Objective) 1

1.2 論文架構 (Thesis Organization) 4

第二章、文獻回顧 (Literature Review) 5

2.1 可靠度分析法 (Reliability Analysis) 5

2.1.1 一階二次可靠度方法 (First Order Second Moment) 5

2.1.2 進階一階二次可靠度方法 (Advanced First Order Second Moment) . . . 6

2.1.3 蒙地卡羅法 (Monte Carlo Simulation) 8

2.1.4 討論 (Discussion) 9

2.2 EGO演算法 (Efficient Global Optimization) 11

2.2.1 貝氏分析 (Bayesian Analysis) 11

2.2.2 演算流程 (Algorithm Flowchart) 12

2.2.3 測試範例 (Demonstration Example) 13

第三章、Kriging模型 (Kriging Model) 16

3.1 基本概念 (Basic Concept) 17

3.2 公式推導 (Kriging Basics Formula Derivation) 21

3.3 參數選擇 (Parameter Estimation) 26

3.3.1 變異圖分析 (Variogram) 26

3.3.2 最大似然推估法 (Maximal Likelihood Estimation) 39

3.4 問題與討論 (Problem and Discussion) 42

第四章、取樣指標 (Infill Sampling Critera) 47

4.1 常用取樣指標 (Common Infill Sampling Criterion) 47

4.1.1 Kushner 47

4.1.2 Expected Improvement 49

4.1.3 Kriging Estimator 51

4.1.4 Maximum Varianve 52

4.1.5 Probability of Feasibility 54

4.1.6 Tunneling 57

4.2 取樣指標分類 (Classification of ISC) 59

4.3 聯合極限狀態取樣指標 (Sampling on Joint Limit State) 61

第五章、演算策略 (Algorithm Strategy) 69

5.1 最佳化 (Optimization) 69

5.2 最大可靠度問題 (Maximal Reliability Problem) 72

第六章、範例 (Example) 74

6.1 凹形可行解空間問題 (Concave Feasible Region Problem) 74

6.2 可靠度最佳化測試範例 (RBDO Test Problem) 77

6.3 調諧減震器問題 (Tuned Vibration Absorber Problem) 83

第七章、結論 (Conclusions) 91

7.1 研究貢獻 (Contribution) 91

7.2 建議 (Suggestion) 92

參考文獻 95

自傳 101

附錄一：中英對照表 102

表　目　錄

2.1 Gomez測試範例在不同演算法下之結果比較表 15

2.2 newBranin測試範例在不同演算法下之結果比較表 15

3.1 似然函數最佳化的結果 43

3.2 樣本空間相關函數矩陣之行列式值 44

4.1 最佳值搜索類型取樣指標表 59

4.2 減少誤差類型取樣指標表 59

4.3 特殊目的類型取樣指標表 60

4.4 兩次取樣RMS表 62

4.5 蒙地卡羅法測試結果表 64

5.1 無拘束條件演算策略之終止條件 70

5.2 有拘束條件演算策略之終止條件 71

5.3 可靠度的問題演算策略之終止條件 73

6.1 凹形可行解空間問題最大可靠度最佳化結果表 75

6.2 指標代號表 76

6.3 RBDO測試範例之最大可靠度最佳化結果表 82

6.4 減震器工程範例破壞機率計算結果表 89

圖　目　錄

1.1 可行解空間在不同可靠度要求下的變化圖 2

2.1 FOSM之誤差示意圖 6

2.2 AFOSM之誤差示意圖 8

2.3 Sasena et al.所使用的EGO演算法流程圖 12

2.4 EGO演算法測試範例圖 14

2.5 EGO演算法測試範例過程圖 14

3.1 取樣與預測位置示意圖 17

3.2 高斯歷程示意圖 18

3.3 取樣與距離示意圖 27

3.4 靠近原點之兩個說明範例樣本圖 28

3.5 距離區域範圍示意圖 29

3.6 Camelback實驗變異圖圖 30

3.7 靠近原點之兩個說明範例樣本圖 31

3.8 靠近原點處之兩個說明範例實驗變異圖 31

3.9 巨觀趨勢之兩個說明範例樣本圖 32

3.10 巨觀趨勢之兩個說明範例實驗變異圖 33

3.11 平穩與非平穩實驗變異圖示意圖 34

3.12 平穩型實驗變異圖名詞示意圖 35

3.13 實驗變異圖與理論變異圖之擬合 38

3.14 最小變異圖誤差擬合結果函數比較圖 38

3.15 最大似然推估法擬合結果函數比較圖 41

3.16 一維測試函數樣本示意圖 42

3.17 不同樣本數之似然函數與θ對照圖 44

3.18 Warenw et al.說明自然對數似然函數具有多峰值和長山脊圖 45

4.1 $\epsilon = 0.1 \times f_{min}$之Kushner指標圖 48

4.2 $\epsilon = 0.01 \times f_{min}$之Kushner指標圖 48

4.3 $g = 1$之改善期望值指標圖 50

4.4 $g = 5$之改善期望值指標圖 51

4.5 Maximum Variance指標圖 52

4.6 利用最大變異數指標修正函數過程圖 53

4.7 取樣與$g = 0$位置圖 55

4.8 可行解機率指標圖 55

4.9 取樣與$g = 0$位置以及新增取樣位置圖 56

4.10 新增取樣後之可行解機率指標圖 57

4.11 新增取樣後之隧道指標圖 58

4.12 Camelback函數與其可行解區域位置圖 62

4.13 第一次取樣結果之Kriging模型與其可行解區域位置圖 63

4.14 第二次取樣結果之Kriging模型與其可行解區域位置圖 63

4.15 範例極限狀態圖 66

4.16 極限狀態取樣指標修正Camelback過程圖 67

4.17 聯合極限狀態取樣指標修正Baseball Function過程圖 68

5.1 EGO演算流程圖 69

5.2 無拘束條件演算策略圖 70

5.3 有拘束條件演算策略圖 71

5.4 可靠度問題演算策略圖 72

5.5 極限狀態取樣演算策略圖 72

6.1 DIRECT演算法之凹形可行解空間問題最大可靠度位置示意圖 75

6.2 可靠度最佳化測試問題圖 77

6.3 RBDO範例之21個樣本極限狀態圖 78

6.4 RBDO範例之30個樣本極限狀態圖 79

6.5 RBDO範例之40個樣本極限狀態圖 80

6.6 RBDO範例之50個樣本極限狀態圖 81

6.7 Zou et al.所使用的調諧減震器示意圖 83

6.8 減震器工程範例問題示意圖 84

6.9 減震器工程範例極限狀態修正過程圖(21、50) 85

6.10 減震器工程範例極限狀態修正過程圖(75、100) 86

6.11 減震器工程範例極限狀態修正過程圖(125、150) 87

6.12 減震器工程範例極限狀態修正過程圖(175、200) 88

6.13 Zou et al.之減震器工程範例的結果圖 89

7.1 Lee et al.所使用的煞車組件示意圖 92

7.2 Camelback函數的Indicator Kriging圖 93

符 號 說 明

$\mathcal{F}$	可行解空間
f	目標函數
$\hat{f}$	目標函數反應曲面預測值
$f_{\mathbf{X}}$	機率密度函數
$f_{\min}$	目標函數取樣中之最佳值
G	拘束條件序率過程函數
$\mathbf{g}$	非等式拘束條件
$\mathbf{g}_{\alpha}$	拘束條件取樣值
$\hat{\mathbf{g}}$	拘束條件反應曲面預測值
h	點與點之間的歐幾里德距離(Euclidean Distance)
I_g	破壞指標函數
L	似然方程式(Likelihood Function)
L_{ν}	拉格朗日方程式(Largrange Eqaution)
$P_{\mathrm{r}}[\cdot]$	$[\cdot]$機率
P_{f}	破壞機率
R	空間相關函數(Spatial Correlation Function)
R_{e}	可靠度
$\mathbf{R}_{\alpha}$	樣本空間相關函數矩陣
$\mathbf{r}$	預測值空間相關函數矩陣
$\mathbf{u}$	標準空間對應之設計變數

$\mathbf{x}$	設計變數
$\mathbf{x}_{\mathrm{f}}$	可行解樣本變數值
$\mathbf{x}_{\mathrm{LB}}$	樣本空間下限
$\mathbf{x}_{\mathrm{UB}}$	樣本空間上限
y	回歸方程式
Z	高斯序率歷程之偏移量
β	回歸方程式係數
β_{r}	可靠度指標(Reliability Index)
$\hat{\beta}$	Kriging 模型母體預估平均值
ϕ	標準高斯機率密度函數
γ	變異圖(Variogram)
γ_{t}	理論變異圖(Theoretical Variogram)
γ_{α}	實驗變異圖(Experimental Variogram)
λ_{α}	Kriging model權重
μ	平均值
ν	拉格朗日乘子(Largrange Multiplier)
Φ	標準高斯累積分佈函數
σ	標準差
ISC	取樣指標(Infill Sampling Criterion)
RMS	方均根誤差 (Root Mean Square)
$Cov[\,]$	共變異數
$E[\,]$	期望值
$Var[\,]$	變異數

第 一 章　緒論 (Introduction)

1.1　研究動機與目的 (Motivation and Objective)

常見的最佳化問題可以用下列方程式(1.1)來表示，其中$\mathbf{x}$爲設計變數，f代表目標函數，$\mathbf{g}$代表非等式拘束條件，$\mathbf{h}$代表等式拘束條件，必須注意到此處的目標函數只有一個，如果有多個目標函數則稱之爲多目標最佳化(Multi-Objective Optimization)，而在本論文中如果沒有註明，則所敘述的最佳化問題都是單目標最佳化問題。在設計空間中滿足拘束條件 $\mathbf{g}(\mathbf{x}) \leq 0$ 以及 $\mathbf{h}(\mathbf{x}) = 0$ 的設計值，我們稱之爲可行解(Feasible Solution)，由可行解所組成的空間，我們稱之爲可行解空間(Feasible Region)，以$\mathcal{F}$表示。

$$
\begin{aligned}
\min_{\mathbf{x}} \quad & f(\mathbf{x}) \\
\text{s.t.} \quad & \mathbf{g}(\mathbf{x}) \leq 0 \\
& \mathbf{h}(\mathbf{x}) = 0 \\
& \mathbf{x}^{\mathrm{LB}} \leq \mathbf{x} \leq \mathbf{x}^{\mathrm{UB}}
\end{aligned}
\tag{1.1}
$$

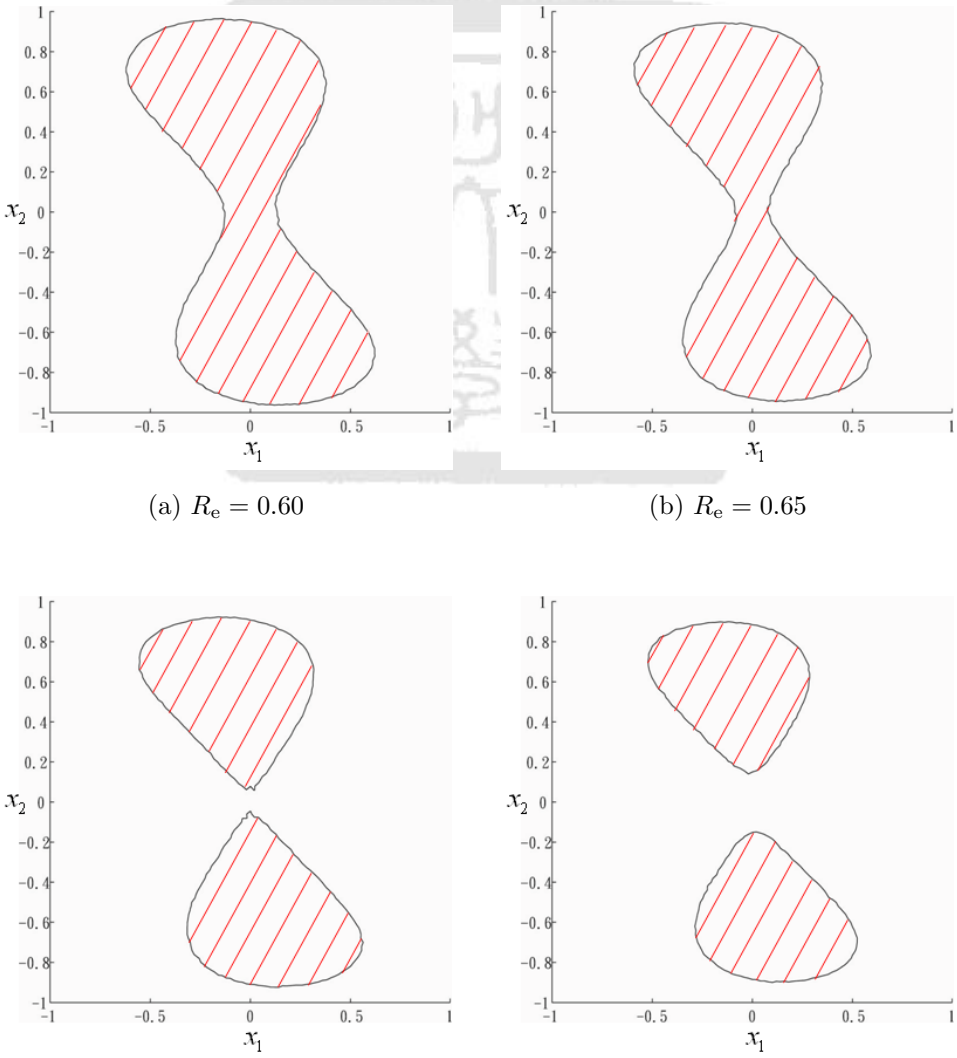
然而在一般的最佳化設計問題中，並沒有考慮不確定因素(Uncertainty)對於問題的影響，但是在一個眞實情形的設計中，不確定因素卻是一個無法避免的現象，而在考慮不確定因素的最佳化設計問題裡，時常會以機率的形式來表示其拘束條件，因此等式限制式並不存在，而這種問題通常可稱之爲可靠度最佳化設計(Reliability-Based Design Optimization)，簡稱爲RBDO [1]。在考慮不確定因素後，RBDO將方程式(1.1)改爲方程式(1.2)，其中假設在此方程式的隨機變數$\mathbf{X}$是一個高斯分佈，$\boldsymbol{\mu}_{\mathbf{x}}$爲設計變數，而$P_{\mathrm{f}}$是違反拘束條件的機率上限，由於違反拘束條件的情形我們稱之爲破壞，因此P_{f}也可以稱之爲破壞機率，所謂的可靠度機率即爲$R_{\mathrm{e}} = 1 - P_{\mathrm{f}}$，而滿足方程式(1.2)之拘束條件的可行解空間，以$\mathcal{F}^{\mathrm{P}}$表示。

$$
\begin{aligned}
\min_{\boldsymbol{\mu}_{\mathbf{x}}} \quad & f(\boldsymbol{\mu}_{\mathbf{x}}) \\
\text{s.t.} \quad & P_{\mathrm{r}}\left[\mathbf{g}(\mathbf{X}) > 0\right] \leq P_{\mathrm{f}} \\
& \boldsymbol{\mu}_{\mathbf{x}}^{\mathrm{LB}} \leq \boldsymbol{\mu}_{\mathbf{x}} \leq \boldsymbol{\mu}_{\mathbf{x}}^{\mathrm{UB}} \\
& \mathbf{X} \backsim N(\boldsymbol{\mu}_{\mathbf{x}}, \boldsymbol{\sigma}_{\mathbf{x}}^2)
\end{aligned}
\tag{1.2}
$$

由方程式(1.2)的拘束條件中，我們可以觀察到，當破壞機率P_{f}的設定值越小，也就是說要求的可靠度R_{e}越高時，則滿足拘束條件的可行解空間$\mathcal{F}^{\mathrm{P}}$將會越小，以一個範例說明，如下列方程式(1.3)的Camelback函數 [2]，考量方程式(1.4)的隨機變數分佈，在不同的破壞機率的情況下，其可行解空間變化情形如圖1.1。

$$g(\mathbf{x}) = -0.1 + \left(4 - 2.1x_1^2 + \frac{x_1^4}{3}\right)x_1^2 + x_1x_2 + \left(-4 + 4x_2^2\right)x_2^2 \leq 0 \tag{1.3}$$

$$\mathbf{X} \backsim \left[N(\mu_{x_1}, 0.2^2), N(\mu_{x_2}, 0.15^2)\right] \tag{1.4}$$



(a) $R_{\mathrm{e}} = 0.60$ (b) $R_{\mathrm{e}} = 0.65$

(c) $R_{\mathrm{e}} = 0.70$ (d) $R_{\mathrm{e}} = 0.75$

圖 1.1: 可行解空間在不同可靠度要求下的變化圖

由圖1.1可以發現可靠度要求越高的情況下，可行解空間會越變越小，當可靠度要求在$65 \backsim 75\%$之間時，原本的一個可行解空間會分裂成兩個小的可行解空間，甚至很有可能根本不存在任何可行解。因此如果我們在不知道拘束條件特性的情況下，隨意要求一個破壞機率作爲拘束條件，將有可能會造成演算法無法找到可行的答案。

一般的工程問題之可靠度往往依照過去的經驗而訂定，爲了瞭解某工程問題可行解空間內可達到之最大可靠度，我們必須先進行最大可靠度的搜尋，如方程式(1.5)，我們稱此方程式爲最大可靠度最佳化問題。

$$\begin{aligned} \min_{\boldsymbol{\mu}_{\mathbf{x}}, P_{\mathrm{f}}} \quad & P_{\mathrm{f}} \\ \text{s.t.} \quad & P_{\mathrm{r}}\left[\mathbf{g}(\mathbf{X}) > 0\right] \leq P_{\mathrm{f}} \\ & \boldsymbol{\mu}_{\mathbf{x}}^{\mathrm{LB}} \leq \boldsymbol{\mu}_{\mathbf{x}} \leq \boldsymbol{\mu}_{\mathbf{x}}^{\mathrm{UB}} \\ & \mathbf{X} \backsim N(\boldsymbol{\mu}_{\mathbf{x}}, \boldsymbol{\sigma}_{\mathbf{x}}^2) \end{aligned} \tag{1.5}$$

由於在方程式中(1.5)，機率形式的拘束條件必爲有效約束(Active Constraint)，因此可以將方程式(1.5)改爲方程式(1.6)，其中由於有多個拘束條件，對於每個拘束條件來說都會計算出一個破壞機率值，由於我們的最佳化架構是單目標最佳化，所以我們必須選擇這些破壞機率裡最大的一個。

$$\min_{\boldsymbol{\mu}_{\mathrm{X}}} \quad \left\{ \max \left[\hat{\mathbf{P}}_{\mathrm{f}}(\boldsymbol{\mu}_{\mathrm{X}}) \right] \right\} \tag{1.6}$$

對於數學問題來說，由於計算函數值所花費的時間很少，計算的時間幾乎可忽略，所以只要需要再重新調整可靠度後執行最佳化即可以解決此問題；但是如果問題本身是一個耗時的黑盒子函數(Expensive Function)時，此黑盒子函數指的是由電腦模擬或是實際實驗等耗費時間才能得到數據的方法，則會顯得既有的方法不具有效率，甚至是不可行。因此，本論文以最大可靠度問題作爲出發點，並以問題本身爲黑盒子函數的架構下，試圖找出一個適合此問題的最佳化演算法。

1.2 論文架構 (Thesis Organization)

本論文一共分爲七章，其內容架構如下：

- 第一章：緒論
  介紹本論文探討問題之背景，並說明本論文之研究動機與目的，並介紹本論文架構。
- 第二章：文獻回顧
  文獻回顧主要內容分爲可靠度分析法與EGO演算法。
- 第三章：Kriging模型
  介紹Kriging模型，包含簡介、基本概念、公式推導、參數抉擇、問題討論...等。
- 第四章：取樣指標
  簡單介紹幾個文獻中的取樣指標，並將這些指標進行分類，然後針對可靠度問題提出一種新的取樣指標。
- 第五章：演算策略
  介紹本論文程式中，針對不同問題所制定的演算策略，並且敘述一些演算法細節。
- 第六章：範例
  使用修正後的EGO演算法，針對不同種類的問題範例進行演算，在此假設這些範例皆爲黑盒子函數。
- 第七章：結論與建議
  總結本論文之內容，並對此論文提出相關建議，以供未來研究發展之參考。

第 二 章　文獻回顧 (Literature Review)

2.1　可靠度分析法 (Reliability Analysis)

本節主要介紹三種常見的可靠度分析法，分別是一階二次可靠度方法、進階一階二次可靠度方法、蒙地卡羅法，並對這三種方法做一個概略性的比較。

2.1.1　一階二次可靠度方法 (First Order Second Moment)

一階二次可靠度方法簡寫爲FOSM，是一階可靠度方法中最爲簡單的一種，此方法的主要內容是在設計點位置的拘束條件函數，利用泰勒展開式(Tayler Series Expansion)做一次展開，然後利用隨機變數的一次與二次統計矩來計算設計點在標準空間中與極限狀態的距離，此標準空間指的是平均值爲零、標準差爲一的標準高斯分佈空間，然後再利用高斯隨機分佈的特性來計算可靠度。

在一個高斯隨機變數$\mathbf{X} \backsim N(\boldsymbol{\mu}_{\mathbf{x}}, \sigma^2_{\mathbf{x}})$與拘束條件$g$的情況下，我們想計算$g(\mathbf{x})$在平均值$\boldsymbol{\mu}_{\mathbf{x}}$的破壞機率，則先對拘束條件在平均值$\boldsymbol{\mu}_{\mathbf{x}}$做一次泰勒展開如下列方程式(2.1)。

$$g'(\mathbf{x}) = g(\boldsymbol{\mu}_{\mathbf{x}}) + \sum_{i=1}^{D} \frac{\partial g}{\partial x_i}(\mathbf{x}_i - \mu_{\mathbf{x}_i}) \tag{2.1}$$

上式中，D代表著此問題的維度，而$g'(\mathbf{x})$是原本拘束條件的線性近似，根據高斯分佈的特性，一個高斯分佈的隨機變數在經過線性轉換後，其本身仍爲一個高斯分佈，也就是說我們將$\mathbf{X}$帶入$g'(\mathbf{x})$後，$g'(\mathbf{X})$會是一個高斯分佈，其平均值與標準差可由方程式(2.2)計算。

$$\begin{aligned} & g'(\mathbf{X}) \backsim N(\mu_{g'}, \sigma^2_{g'}) \\ & \begin{cases} \mu_{g'} = g(\boldsymbol{\mu}_{\mathbf{x}}) \\ \sigma^2_{g'} = \sum_{i=1}^{D} \left(\frac{\partial g}{\partial x_i}\right)^2 \sigma^2_{\mathbf{x}_i} \end{cases} \end{aligned} \tag{2.2}$$

由於隨機變數是高斯分佈，因此我們可以利用高斯分佈的特性，直接計算$g'(\mathbf{X})$的破

壞機率，如方程式(2.3)，其中β_{r}稱爲可靠度指標(Reliability Index)。

$$\hat{P}_{\mathrm{f}} = P_{\mathrm{r}}\left[g'(\mathbf{X}) > 0\right] = \Phi\left(\frac{\mu_{g'} - 0}{\sigma_{g'}}\right) = \Phi(\beta_{\mathrm{r}}) \tag{2.3}$$

此方法由於只需計算拘束條件的梯度，因此在計算速度上相當快速，然而此方法所計算的對象是拘束條件的一次線性近似方程式，如果拘束條件並非線性函數則此方法將不準確，如下圖2.1，一個一維的非線性拘束條件示意圖。

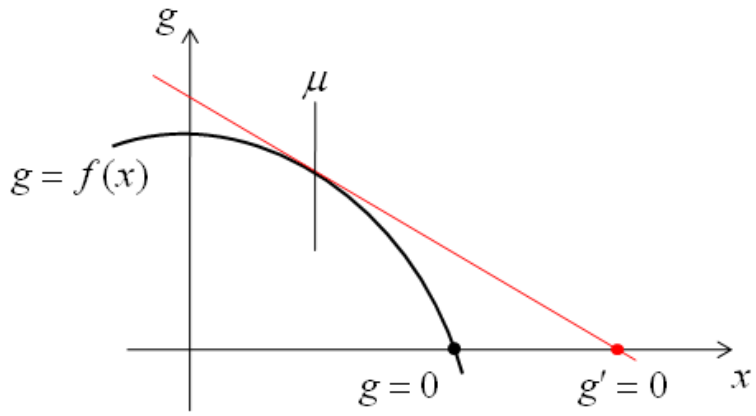


圖 2.1: FOSM之誤差示意圖

上圖中，黑色實線爲眞實的拘束條件，紅色實線爲泰勒展開式在平均值位置的一次展開，由此圖可以發現當拘束條件爲非線性時，隨機變數的平均值位置與眞實函數極限狀態之間的距離，將與FOSM所預測的距離不一樣，而此距離的誤差將會影響可靠度指標的計算；由於FOSM在使用上的局限性，使其後來發展另一種一階二次可靠度法，我們稱之爲進階一階二次可靠度法。

2.1.2 進階一階二次可靠度方法 (Advanced First Order Second Moment)

進階一階二次可靠度法簡寫爲AFOSM，是由Hasofer與Lind在1974年所提出的方法 [3]，因此也被稱爲H-L法，是目前最常見的可靠度分析方法。AFOSM的概念相當簡單，藉由搜索在標準空間中距離設計點最近的拘束條件位置，則兩者之間的距離即爲可靠度指標。

在一個高斯隨機變數X與拘束條件g的情況下，我們想計算在平均值μ_{x}的破壞機率，

則先對設計變數做轉換，將其轉至標準空間中，如方程式(2.4)，其中設計空間指的是$\mathbf{x}$變數的空間，標準空間指的是平均值爲零、標準差爲一的空間，以$\mathbf{u}$代表。

$$\mathbf{u} = \frac{\mathbf{x} - \mu_{\mathbf{x}}}{\sigma_{\mathbf{x}}} \tag{2.4}$$

在標準空間中的原點即爲設計空間中的平均值，由於隨機變數的標準差已經過標準化，因此在標準空間中距離原點越近的地方即代表擁有越大的機率，所以在此空間中相對應的拘束條件上距離原點最近的點，即爲最有可能破壞點(Most Probable Point，MPP)，其搜索距離最近的點，如下列方程式(2.5)所示。

$$\begin{aligned} &\min_{\mathbf{u}} \|\mathbf{u}\| \\ &\text{subject to } \ g(\mathbf{u}) = 0 \end{aligned} \tag{2.5}$$

由上式最佳化過程所找到在設計空間中的最佳值$\|\mathbf{u}^*\|$，即爲可靠度指標β_{r}，因此我們就可以利用此值來計算破壞機率，如方程式(2.6)。

$$\hat{P}_{\mathrm{f}} = \Phi(\beta_{\mathrm{r}}) = \Phi(\|\mathbf{u}^*\|) \tag{2.6}$$

將此方法歸類於一階方法，是因爲在得到標準空間中的MPP後，其計算可靠度的概念是將拘束條件在標準空間中的MPP位置做一階泰勒展開，如圖2.2，因此AFOSM計算破壞機率的準確性，與在標準空間中拘束條件極限狀態的線性程度有關，由於標準空間與變數空間之間是高斯轉換，因此在標準空間中爲線性即代表在變數空間中也爲線性，因此AFOSM的準確性直接與變數空間中拘束條件極限狀態的線性與否相關；另外一個影響AFOSM的準確性是有多個MPP時，由於最佳化的結果只能呈現一個MPP，因此當有多個MPP與原點距離相差不遠時，AFOSM將會低估破壞機率。而在AFOSM的概念下，實際應用在可靠度最佳化演算法中常用演算的方法爲下列兩種：可靠度指標法(Reliability Index Approach) [3,4]與性能測量法(Performance Measure Approach) [5,6]。

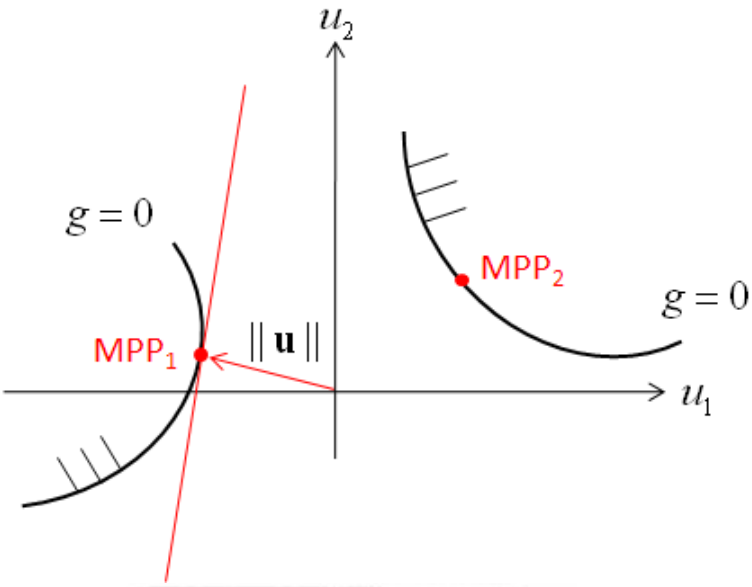


圖 2.2: AFOSM之誤差示意圖

2.1.3 蒙地卡羅法 (Monte Carlo Simulation)

蒙地卡羅法(Monte Carlo Simulation)屬於取樣法(Sampling)的一種，簡稱爲MCS，是一種常見的可靠度分析法，可參考 [7]。實際使用蒙地卡羅法計算可靠度時，是利用電腦大量製造符合隨機變數機率分佈的樣本點，再將這些樣本代入拘束條件來計算可靠度，與一階可靠度方法最大的不同是，它並不限制隨機變數的種類是否爲高斯分佈，由於此方法的概念簡單，因此時常被拿來作爲驗證答案的可靠度準確性的方法。

假設有一個隨機變數X與拘束條件g，蒙地卡羅法計算可靠度的方式，如下列方程式(2.7)所示，其中$\mathbf{x}_i$是由母體X所取出之樣本，總數爲N個樣本，而$I_g(\mathbf{x})$僅代表一種指標來表示樣本點的破壞與否，而最後計算的破壞機率即爲$\hat{P}_\mathrm{f}$。

$$
\begin{aligned}
I_g(\mathbf{x}_i) &= \begin{cases} 1, & \text{if } g(\mathbf{x}_i) > 0 \\ 0, & \text{else} \end{cases} \\
\hat{P}_\mathrm{f}(\mathbf{X}) &= \frac{1}{N}\sum_{i=1}^{N} I_g(\mathbf{x}_i)
\end{aligned}
\tag{2.7}
$$

然而對一個存在於連續空間中的隨機變數**X**而言，其破壞機率P_{f}的實際值應等於方程式(2.8)。

$$P_{\mathrm{f}} = \int_{-\infty}^{\infty} I_g(\mathbf{x}) f_{\mathrm{x}}(\mathbf{x}) d\mathbf{x} \tag{2.8}$$

但是我們永遠無法得到此破壞機率的實際值，根據文獻 [7]，蒙地卡羅法所計算的破壞機率$\hat{P}_{\mathrm{f}}$與實際破壞機率P_{f}之間的誤差，如下列方程式(2.9)，其中N爲樣本數，ε爲計算的破壞機率與實際的破壞機率的誤差百分比。

$$\varepsilon\% = \sqrt{\frac{1 - P_{\mathrm{f}}}{N \times P_{\mathrm{f}}}} \times 200\% \tag{2.9}$$

經由方程式(2.9)的計算，如果實際的破壞機率爲1%，則在取樣10^6個樣本點與95%的信賴區間的情況下，蒙地卡羅計算的破壞機率將與實際的破壞機率相差大約0.02%左右；如果實際的破壞機率爲10%，則在取樣10^6個樣本點與95%的信賴區間的情況下，蒙地卡羅計算的破壞機率將與實際的破壞機率相差大約0.6%左右。

2.1.4 討論 (Discussion)

實際上在文獻中的可靠度分析法並不只有以上介紹的內容，其他方法還有：二階可靠度方法(Second Order Reliability Method) [7]、統計矩法(Moment Method) [8]、重要取點法(Importance Sampling) [9]...等，我們姑且將這些方法分成兩種概念。

1. 第一種概念：利用拘束條件的函數量值來計算可靠度。
   如：進階一階二次可靠度法利用高斯分佈的特性，藉由計算設計點與M.P.P.之間的距離來估算可靠度；統計矩法是計算拘束條件在隨機變數下的四個統計矩來計算可靠度。

2. 第二種概念：直接利用取樣對實際破壞機率P_{f}進行逼近。
   如：蒙地卡羅法、重要取點法...等。

在此兩個概念下的可靠度分析法，以蒙地卡羅法與進階一階二次可靠度法在使用上最爲簡單，因此我們在此列舉此兩個方法的優缺點。

- 蒙地卡羅法的優點有：

  1. 隨機變數不局限於高斯分佈。
  2. 在不預設拘束條件的線性程度下，可以達到較爲準確的結果。

- 蒙地卡羅法的缺點有：

  1. 需要大量的樣本才能有一定的準確性。
  2. 由於隨機取樣的關係，同個設計點重複進行蒙地卡羅法將不會得到一樣的答案，造成演算法進行最佳化的困難。

- 進階一階二次可靠度法的優點有：

  1. 取樣較少。
  2. 相同位置所計算之可靠度計算是一樣的值。
  3. 在拘束條件極限狀態爲線性且不具有多個MPP的情況下，可以達到完全準確的結果。

- 進階一階二次可靠度法的缺點有：

  1. 隨機變數必須是高斯分佈。
  2. 在無法確定拘束條件的形式下，可靠度計算值並不準確。

在本論文程式中，使用者可以選擇一階二次可靠度法、進階一階二次可靠度法或蒙地卡羅法來進行可靠度的計算。然而上述三種方法在計算可靠度時，都必須對原函數進行額外的函數計算，因此在問題本身爲黑盒子函數的情況下，會使得這些可靠度分析法都不夠有效率，甚至是不可行，舉例來說，如果計算一次黑盒子函數需要花費一個小時，使用蒙地卡羅法取10^3個樣本點計算可靠度，那麼光是計算一個設計值的可靠度就必須花費41天。因此，傳統上處理這種黑盒子函數的可靠度問題，往往是將可靠度分析法與反應曲面法結合，以減少計算可靠度的時間，關於此部份可參考第四章聯合極限狀態搜索指標。

2.2 EGO演算法 (Efficient Global Optimization)

Efficient Global Optimization簡稱爲EGO，是一種針對問題本身是黑盒子函數的最佳化演算法，此演算法在1998年由Jones、Schonlau、Welch [10]等人提出，是一個發展歷史並不久遠的演算法，由於此演算法是針對黑盒子函數所設計，因此能夠減少對原黑盒子函數問題的取樣(Sampling)或函數計算(Function Count)是此演算法的主要特色。

EGO演算法是一種使用Kriging模型作爲反應曲面的演算法，而此種使用統計模型作爲反應曲面的演算法，我們稱之爲貝氏分析法，而EGO演算法藉由更新反應曲面的方式來得知全域最佳化(Global Optimum)，是一種二階段(Two-Stage)型的演算法，此二階段的涵義是，演算法在決定下一個計算位置時，是利用反應曲面的資訊進行一次的最佳化過程來決定，關於使用反應曲面的全域最佳化演算法分類，可參考文獻 [11]。

而在1998年之後，EGO演算法幾個比較重要的發展包含了，2001年由Gutmann發展以徑向基底函數(Radial Basis Function)作爲反應曲面的演算法 [12]；2002年由Sasena將EGO延伸至黑盒子函數與純數學問題之間的混合型問題 [13]；2006年由Knowles將 EGO演算法推廣至多目標最佳化問題 [14]。EGO所使用的反應曲面是一種名稱爲Kriging的統計模型，此統計模型在第三章有詳細的介紹，以下針對EGO的細節加以描述。

2.2.1 貝氏分析 (Bayesian Analysis)

貝氏分析(Bayesian Analysis)爲一種使用統計模型的演算法，其歷史可追溯至1960年代。最早的貝氏分析演算法所使用的統計模型是韋納過程(Wiener Process)，也有人稱之爲布朗運動(Brownian Motion)或是馬可夫過程(Markov Process)，主要概念是某位置的函數期望值與變異數，與某位置距離最近的樣本有關，可參考 [15]。而最早將貝氏分析加入最佳化演算的概念是由Kushner於1964年提出 [16]。

最初Kushner提出此概念的主要原因在於梯度型(Gradient-Based)演算法無法克服局限於局部最佳值(Local Optima)的問題，Kushner認爲在演算法決定下一個點的位置時，不要只使用局部的資訊來決定，而是使用所有取樣的資訊來決定，因此在Kushner的方法中，下一個點是反應曲面上擁有最大改善機率的位置，如方程式(2.10)，其中$\hat{f}(x)$是反應曲面上的數值，$f_{\min}$是樣本點中最小的值，ϵ爲使用者自定義數。

$$\max_{x} \quad P_{\mathrm{r}}\left[\hat{f}(x) < f_{\min} - \epsilon\right] \tag{2.10}$$

在演算法中，Kushner藉由調整ϵ的數值來決定演算法的搜尋性質，此方法後來被人稱之爲Kushner方法，而方程式(2.10)後來被稱爲Kushner指標，其內容可以參考本論文第四章，欲瞭解更多有關於貝氏分析演算法的內容，可參考 [17]。

2.2.2 演算流程 (Algorithm Flowchart)

EGO演算法使用Kriging模型，利用貝氏分析演算法的概念，在每一次迭代時都對反應曲面所組成的函數進行一次最佳化，由於此反應曲面所組成的函數是要讓演算法決定下一個取樣位置的指標，因此我們稱此函數爲取樣指標(Infill Sampling Criterion，ISC)，然而取樣指標並不一定只限於與目標函數有關，關於取樣指標的內容可參考本論文第四章，根據文獻 [18]中，一個EGO演算法的流程如下圖2.3所示。

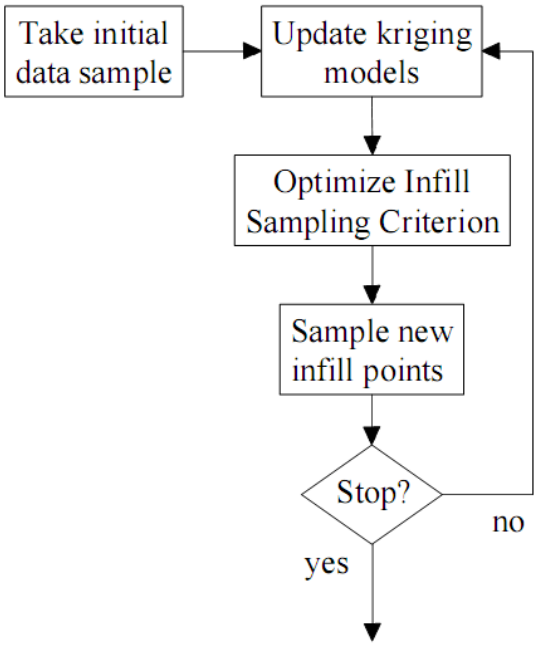


圖 2.3: Sasena et al.所使用的EGO演算法流程圖

如上圖，EGO演算法主要分爲五個步驟，分別是

1. 對問題的變數空間進行初始取樣(Initial Sampling)，取樣方法可以用拉丁方格法(Latin-Hypercube)、田口式直交表法(Taguchi Orthogonal Arrays)、全因子實驗法(Full-Factorial)等。
2. 根據取樣，建立Kriging模型。
3. 根據Kriging模型的資訊，對取樣指標進行最佳化，可以在同一次迭代中，針對多種取樣指標進行最佳化。
4. 根據第三步驟之結果，對眞實函數進行取樣。
5. 當取樣結果之函數改善量值過小，則演算結束，否則回到第二步驟。

由流程圖的步驟中可以發現，EGO的主要內容在於Kriging模型與取樣指標，必須注意到在建立Kriging模型時，需要一個最佳化的過程，取樣指標也需要一個最佳化的過程，所以在EGO演算法中的一次迭代就包含了兩個最佳化的過程，因此若是原問題並非黑盒子函數，則EGO演算法並沒有太大的優勢。

2.2.3 測試範例 (Demonstration Example)

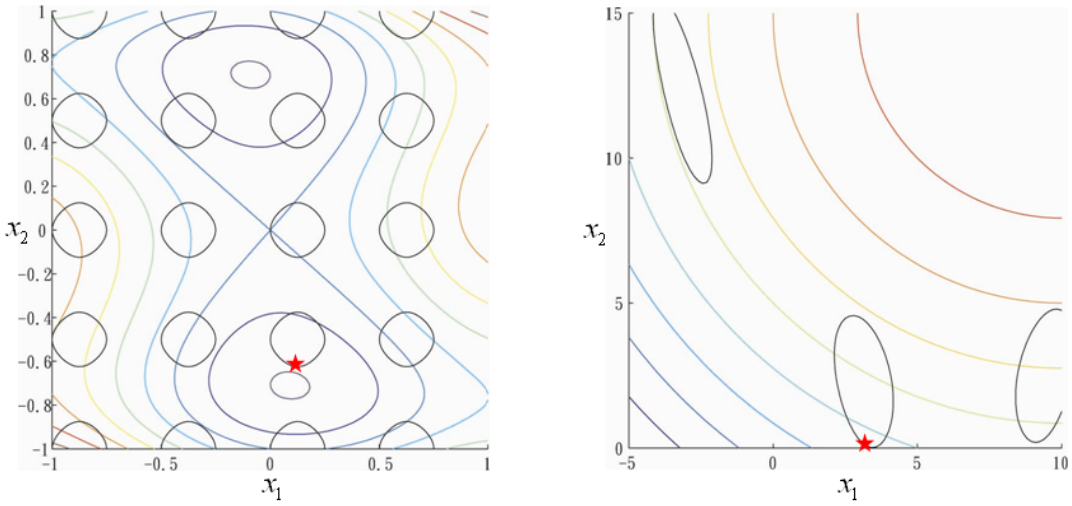
以下利用兩個範例來說明EGO演算法與其他全域最佳化演算法的差異。

Gomez測試範例 [2,19]：

$$\begin{aligned} \min_{\mathbf{x}} \quad & f(\mathbf{x}) = \left(4 - 2.1x_1^2 + \frac{x_1^4}{3}\right)x_1^2 + x_1x_2 + \left(-4 + 4x_2^2\right)x_2^2 \\ \text{s.t.} \quad & g(\mathbf{x}) = -\sin(4\pi x_1) + 2\sin^2(2\pi x_2) \leq 0 \end{aligned}$$

newBranin測試範例 [20,21]：

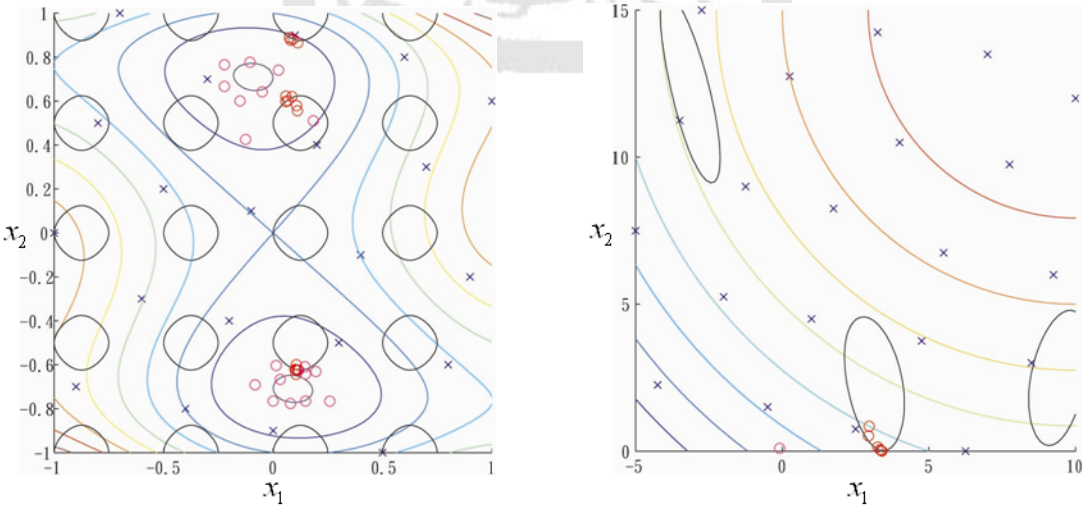
$$\begin{aligned} \min_{\mathbf{x}} \quad & f(\mathbf{x}) = -(x_1 - 10)^2 - (x_2 - 15)^2 \\ \text{s.t.} \quad & g(\mathbf{x}) = \left(x_2 - \frac{5.1}{4\pi^2}x_1^2 + \frac{5}{\pi}x_1 - 6\right)^2 + 10\left(1 - \frac{1}{8\pi}\right)\cos x_1 + 10 \leq 0 \end{aligned}$$



(a) Gomez (b) newBranin

圖 2.4: EGO**演算法測試範例圖**

此兩個範例如圖2.4**，圖中黑色實線爲拘束條件等於零的位置，我們稱此位置爲極限狀態**(Limit State)**，等高線爲目標函數的走向，星號爲問題的最佳解。**EGO**演算過程如圖**2.5**所示，圖中**x**爲初始樣本，**o**爲**EGO**演算過程的取樣。**



(a) Gomez (b) newBranin

圖 2.5: EGO**演算法測試範例過程圖**

將EGO演算法的結果與MATLAB程式之基因演算法(Genetic Algorithm，GA)、密西根大學Whitehead等人所撰寫之模擬退火演算法(Simulated Annealing Algorithm，SA)、密西根大學Sasena所撰寫之DIRECT演算法等最佳化演算法進行比較，如表2.1、表2.2。

表 2.1: Gomez測試範例在不同演算法下之結果比較表

	函數計算次數	最佳值
EGO	55	-0.9710
DIRECT	173	-0.9678
GA	3709	-0.9709
SA	4241	-0.9707

表 2.2: newBranin測試範例在不同演算法下之結果比較表

	函數計算次數	最佳值
EGO	42	-268.36
DIRECT	119	-268.78
GA	5201	-268.76
SA	5371	-268.51

由上述兩個測試範例可以看出，EGO演算法在函數的總計算次數上，都遠比其他最佳化演算法來的少，這也是此演算法最大的特點，然而各種演算法都有擅長之處，在此我們並不是想比較演算法的優劣，誠如之前所提到，如果問題本身並非需要較長時間計算的黑盒子函數，則EGO演算法在演算時間上將不具有優勢。

第 三 章　Kriging模型 (Kriging Model)

Kriging模型最原始的基本概念，源自於南非的礦冶工程師Krige對於南非礦脈分佈的探查研究 [22]，而後由法國數學家Matheron根據Krige的研究，提出一套有別於傳統統計方法的估算理論來作爲地理統計學(Geostatistics)的基礎，地理統計又稱爲空間統計(Spatial Statistics)，於1950年代初期開始發展形成一門新的統計學分支，而Kriging模型是地理統計學中，一種用來推估量值的方法。

在1940年代，地理學家在南非進行礦場資源估計與評價時，發現傳統統計學沒有考量樣本區塊體積的大小與空間變化的特徵，因此在進行估計與評價上存在許多不能克服的問題。後來統計學者發現這些變量並不是單純的隨機變量，因此不能用簡單的統計方法進行估計與評價。到了1950年代初期，南非的礦冶工程師Krige與Sichel研究南非金礦儲量計算的問題，提出了用樣本空間位置和相關程度來估計區塊位置及金礦儲存量，並將其估計誤差最小化的金礦儲存量計算概念，是爲Kriging模型的雛型。

1963年，Matheron出版Principles of Geostatistics，提出地理統計學的概念，定義地理統計學是一種隨機函數(Random Function)在推估自然現象的應用，並在1970年提出區域化變數理論，探討自然現象在空間上分佈的相關性，爲了紀念Krige的貢獻，Matheron將此推估方法方法命名爲Kriging。1970年代以後，大量地理統計研究與應用書籍相繼出版，至此，地理統計理論已被廣泛應用於環境工程、生態、水文、遙測、人文等空間資料分析，由於此方法最初是因應地理學領域(如採礦學、地質學)的需求而發展與應用，所以被稱爲地理統計學，而Matheron也因此被人稱爲地理統計學的奠基者，其生平可以參考 [23]，欲瞭解更多有關於Kriging的起源與歷史，可以參考 [24]。

總結上述，我們可以簡單的說，Kriging模型的主要內容，在於探討如何在有限的觀察數據下，建立一個在特定範圍內某種自然現象變異程度的預測模型。

到了1980年代末期，Kriging模型開始被應用到電腦實驗模擬(Computer Experiments) [25,26]，作爲一種類似於反應曲面法(Response Surface Mathodology)的工具，而這種結合實驗設計與Kriging模型來預測輸入與輸出之間的關係之方法，後來被人稱爲DACE(Design and Analysis of Computer Experiments)，至此，Kriging模型開始被廣泛應用在各種工程領域問題上，如直昇機測試實驗 [27]、熱流通道散熱孔設計 [28]、剛性薄板設計 [29]、儲水壓力桶設計 [30]...等。然而DACE與地理統計學對於Kriging模型的使用上稍有不同，最大

的不同點是在Kriging 模型的參數抉擇，兩者有著截然不同的作法，此部份會在本章第四節做比較。

簡單的說DACE所使用的Kriging模型是地理統計學中的其中一部分，而兩種概念在描述上可以分別參考 [31,32] ，本論文主要會以地理統計學的角度來講解Kriging模型的基本概念。

3.1 基本概念 (Basic Concept)

欲觀察某個區域內的一種自然現象之趨勢或分佈，如海拔高度、空氣密度、溫度、濕度等，直觀的作法是在不同的位置取樣量測此自然現象，使我們能對這個自然現象在空間中的分佈有初步瞭解，然後再根據這些取樣所收集到的數據，來預測在這個區域內的一個沒有經過取樣的位置之預測值，如圖3.1，o爲取樣點、x爲欲預測的位置。

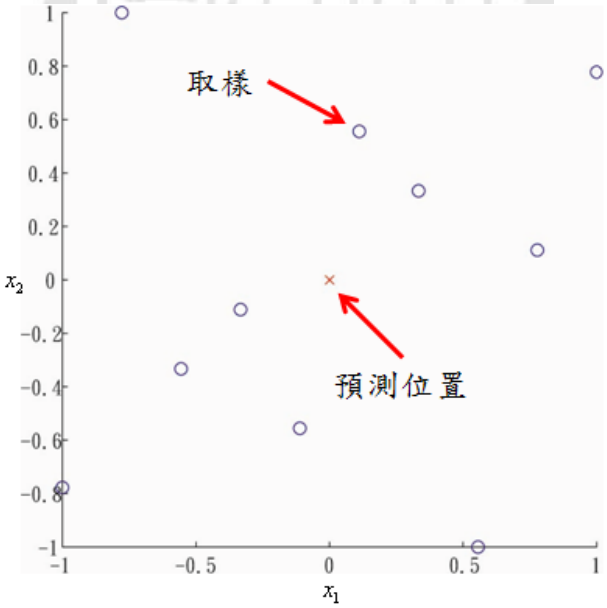


圖 3.1: 取樣與預測位置示意圖

如果由n_s個取樣點所得到的量測值爲g_α，未取樣位置的函數預測值爲$\hat{g}$，Kriging模型的概念是此預測值是所有量測值的線性組合，如方程式(3.1)，其中λ_α爲一組線性組合的權重，而不同的位置會有不同的權重，因此對於Kriging模型來說，只要決定了這組權重即可決定取樣與預測之間的關係。

$$\hat{g}(\mathbf{x}) = \sum_{\alpha=1}^{n_s} \lambda_\alpha(\mathbf{x}) g_\alpha \tag{3.1}$$

然而可以發現一個事實是即使是在相同的位置，重複的量測此自然現象所得到的量測值，並不一定都會是同一個量值，往往可以發現此量測值是一個隨機函數，我們稱這個現象叫做序率歷程(Stochastic Process)，如果此序率函數是一個高斯分佈，我們則可以說這個自然現象在空間中的不同位置是一個高斯歷程(Gaussian Process)，如圖3.2，粗實線代表平均值，細實線代表在某個位置的分佈，必須注意到Kriging模型並不假設其本身爲某種形式的序率歷程，然而在此我們將Kriging模型假設爲高斯歷程來進行說明，而在本論文中所有的Kriging模型都將以高斯歷程來代表。

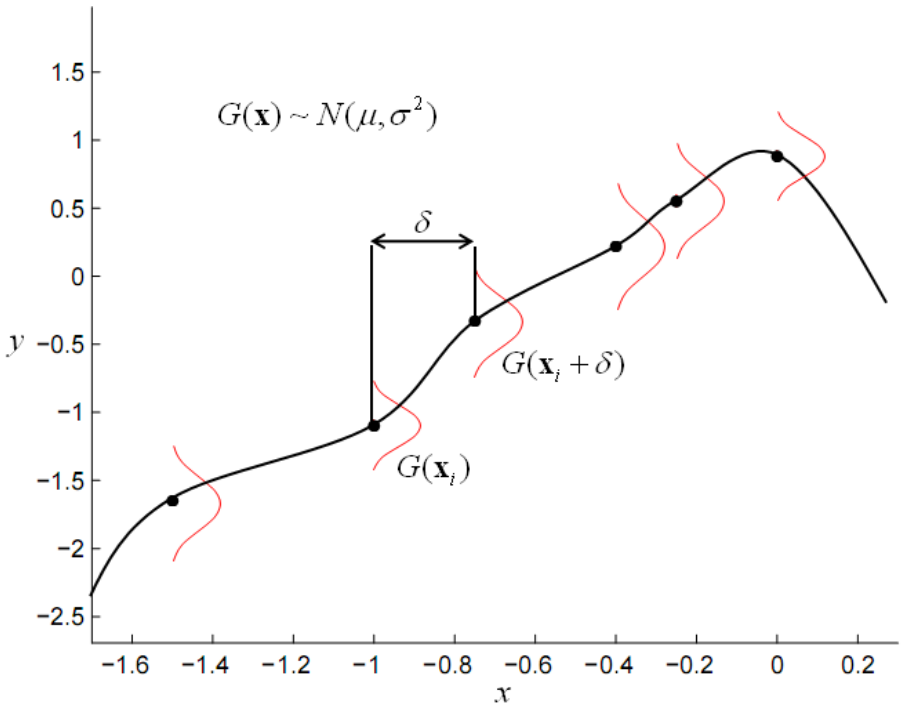


圖 3.2: 高斯歷程示意圖

因此針對一個高斯歷程的函數$G(\mathbf{x})$，我們可以將其分解改寫成定率(Deterministic)與序率(Stochastic)的兩個部份，如方程式(3.2)，其中定率部份的$y_j(\mathbf{x})$爲回歸函數、β_j爲回歸係數，序率部份的$Z(\mathbf{x})$是一個高斯分佈，也可以稱之爲偏差值(Departure)，而此高斯分佈是一個以平均值爲零、標準差爲σ_z的隨機分佈。

$$G(\mathbf{x}) = \sum_{j=1}^{k} \beta_j y_j(\mathbf{x}) + Z(\mathbf{x}) \tag{3.2}$$

像這樣一個存在於空間中的序率歷程，其本身是否存在著一些統計上的特性？如圖3.2，簡單的說如果兩個相距δ的隨機函數分別爲$G(\mathbf{x}_i)$、$G(\mathbf{x}_i+\delta)$，其對應的偏差值爲$Z(\mathbf{x}_i)$與$Z(\mathbf{x}_i+\delta)$，假設δ是一個相當小的值，Kriging理論認爲如果在$\mathbf{x}_i$位置的偏差值很大，那麼就可以很合理的認爲在$\mathbf{x}_i+\delta$位置的偏差值也會很大，也就是說一個序率歷程中的兩個不同位置之偏差值的關係，會與兩者之間的距離有關。因此，在經過大量統計數據上的觀察後，統計學家提出了兩種關於隨機函數的基本假設，分別是平穩假設(Stationay Hypothesis)與本質假設(Intrinsic Hypothesis)，分別如下：

- 平穩假設：完整的敘述應該是二階平穩假設(Second-Order Stationary Hypothesis)，又稱爲廣義性平穩假設(Wide-Sense Stationary)，當一個隨機函數滿足下列三個條件則稱爲滿足平穩假設。

  1. 隨機函數在不同位置的期望值不會因爲位置的不同而改變。

  $$E[G(\mathbf{x})] = m \tag{3.3}$$

  2. 隨機函數在不同位置的變異數不會因爲位置的不同而改變。

  $$Var[G(\mathbf{x})] = E[G(\mathbf{x})^2] = E[Z(\mathbf{x})^2] = \sigma_z^2 \tag{3.4}$$

  3. 兩個不同位置的隨機函數，其共變異數只與兩者之間的相對距離有關，與其分別所在位置無關。

  $$\begin{aligned} Cov(G(\mathbf{x}_i), G(\mathbf{x}_j)) &= Cov(Z(\mathbf{x}_i), Z(\mathbf{x}_j)) \\ &= E[(Z(\mathbf{x}_i)-m)(Z(\mathbf{x}_j)-m)] = C_1(h) \end{aligned} \tag{3.5}$$

  其中$h = \|\mathbf{x}_i - \mathbf{x}_j\|$，$C_1(h)$爲一個與距離有關之函數。

- 本質假設：當隨機函數滿足下列兩個條件則稱爲滿足本質假設。

  1. 兩個不同位置的隨機變數，其之間差值的期望值爲兩個位置之間距離的函數，與其分別所在位置無關。

  $$E[G(\mathbf{x}_i) - G(\mathbf{x}_j)] = C_1(h) \tag{3.6}$$

2. 兩個不同位置的隨機變數，其之間差值的變異數爲兩個位置之間距離的函數，與其分別所在位置無關。

$$Var[G(\mathbf{x}_i) - G(\mathbf{x}_j)] = E[(G(\mathbf{x}_i) - G(\mathbf{x}_j))^2] = C_2(h) \tag{3.7}$$

其中$h = \|\mathbf{x}_i - \mathbf{x}_j\|$，$C_1(h)$與$C_2(h)$僅代表兩個與距離有關之函數。

必須注意到，一個滿足平穩假設的隨機函數，通常也會滿足本質假設，但是一個滿足本質假設的函數，並不一定會滿足平穩假設，而這些假設是一種統計數據上觀察的結果，其本身並不代表一種自然定律，因此並非所有自然現象會滿足上述假設，然而我們可以根據上述假設所得到的統計特性，來對Kriging模型的方程式進行推導，其推導過程可見下一節。總結上述，Kriging模型的基本概念主要可以分爲下列兩點：

1. 高斯歷程中兩個不同位置的偏差值，其共變異數與兩者相對位置有關。
2. 預測值是量測值的線性組合。

3.2 公式推導 (Kriging Basics Formula Derivation)

此部份公式推導可參考文獻 [33]。假設存在於空間中的自然現象表示成方程式(3.2)，其中$y_j(\mathbf{x})$爲回歸函數、β_j爲回歸係數，如果我們用k個方程式來組成這個回歸函數，則我們可以表示成下列$k \times 1$的矩陣，以及相對應的回歸係數矩陣，並且可以將方程式(3.2)改寫成方程式(3.8)。

$$\mathbf{y}(\mathbf{x}) = \left[\ y_1(\mathbf{x}), y_2(\mathbf{x}), \ldots, y_k(\mathbf{x})\ \right]^T$$

$$\boldsymbol{\beta} = \left[\ \beta_1, \beta_2, \ldots, \beta_k\ \right]^T$$

$$G(\mathbf{x}) = \mathbf{y}(\mathbf{x})^T \boldsymbol{\beta} + Z(\mathbf{x}) \tag{3.8}$$

假設取n_s個樣本點，分別爲$\{\mathbf{x}_\alpha : \alpha = 1, \ldots, n_s\}$，所得到的量測值分別爲$\{g_\alpha : \alpha = 1, \ldots, n_s\}$，將這些樣本結果代入方程式(3.8)可得到下列方程式(3.9)，根據假設，我們將此關係式視爲一個實際存在的函數。

$$\begin{pmatrix} g_1 \\ g_2 \\ \vdots \\ g_{n_s} \end{pmatrix} = \begin{pmatrix} y_1(\mathbf{x}_1) & y_2(\mathbf{x}_2) & \cdots & y_k(\mathbf{x}_{n_s}) \\ y_1(\mathbf{x}_1) & & & \vdots \\ \vdots & & & \\ y_1(\mathbf{x}_1) & \cdots & & y_k(\mathbf{x}_{n_s}) \end{pmatrix} \begin{pmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_k \end{pmatrix} + \begin{pmatrix} Z(\mathbf{x}_1) \\ Z(\mathbf{x}_2) \\ \vdots \\ Z(\mathbf{x}_{n_s}) \end{pmatrix}$$

$$\mathbf{g}_\alpha = \mathbf{y}_\alpha \boldsymbol{\beta} + \mathbf{Z} \tag{3.9}$$

根據方程式(3.1)，Kriging模型是樣本點測量值的線性組合，我們稱之爲Kriging Estimator，如方程式(3.10)

$$\begin{aligned} \hat{g}(\mathbf{x}) &= \boldsymbol{\lambda}_\alpha(\mathbf{x})^T \mathbf{g}_\alpha \\ &= \boldsymbol{\lambda}_\alpha(\mathbf{x})^T \mathbf{y}_\alpha \boldsymbol{\beta} + \boldsymbol{\lambda}_\alpha(\mathbf{x})^T \mathbf{Z} \end{aligned} \tag{3.10}$$

在推導Kriging模型的過程中，所使用的概念是最佳無偏估線性預測法(Best Linear Unbiased Predicto，BLUP)，因此推導過程主要分爲兩個步驟，分別是無偏差(Unbiased)與最小均方誤差(Mean Square Error)兩個部份；第一個步驟是令Kriging模型具有無偏差的性質，方法是將眞實函數方程式(3.8)與Kriging Estimator方程式(3.10)的期望值相等，可得方程式(3.11)作爲無偏差拘束條件，即當線性組合的權重滿足此方程式時，Kriging模型即可具有無偏差性質的預測值。

$$
\begin{aligned}
E[\hat{g}(\mathbf{x})] &= E[G(\mathbf{x})] \\
E[\boldsymbol{\lambda}_\alpha(\mathbf{x})^T\mathbf{y}_\alpha\boldsymbol{\beta} + \boldsymbol{\lambda}_\alpha(\mathbf{x})^T\mathbf{Z}] &= E[\mathbf{y}(\mathbf{x})^T\boldsymbol{\beta} + Z(\mathbf{x})] \\
\boldsymbol{\lambda}_\alpha(\mathbf{x})^T\mathbf{y}_\alpha\boldsymbol{\beta} &= \mathbf{y}(\mathbf{x})^T\boldsymbol{\beta} \\
\mathbf{y}_\alpha^T\boldsymbol{\lambda}_\alpha(\mathbf{x}) &= \mathbf{y}(\mathbf{x})
\end{aligned} \tag{3.11}
$$

在進行接下來的推導之前，我們先從新檢視由平穩假設所得到的方程式(3.5)，將兩個隨機函數的共變異數方程式定義爲以下方程式(3.12)，其中$R(\mathbf{x}_i, \mathbf{x}_j)$稱爲空間相關函數(Spatial Correlation Function)，爲了推導方便，在此先將此方程式視爲一個已知的函數，其詳細內容將在下一節說明。

$$
E[Z(\mathbf{x}_i)Z(\mathbf{x}_j)] = Cov(Z(\mathbf{x}_i), Z(\mathbf{x}_j)) = \sigma_z^2 R(\mathbf{x}_i, \mathbf{x}_j) \tag{3.12}
$$

第二步驟是要最小化均方誤差，將Kriging Estimator與眞實函數之間的均方誤差列出，爲了簡化符號，我們將$\boldsymbol{\lambda}_\alpha(\mathbf{x})$省略爲$\boldsymbol{\lambda}_\alpha$，如下

$$
\begin{aligned}
E[\hat{g}(\mathbf{x}) - G(\mathbf{x})]^2 =& E[\boldsymbol{\lambda}_\alpha^T\mathbf{g}_\alpha - G(\mathbf{x})]^2 \\
=& E[\boldsymbol{\lambda}_\alpha^T\mathbf{g}_\alpha\mathbf{g}_\alpha^T\lambda_\alpha - 2\boldsymbol{\lambda}_\alpha^T\mathbf{g}_\alpha G(\mathbf{x}) + G(\mathbf{x})^2] \\
=& E[\boldsymbol{\lambda}_\alpha^T(\mathbf{y}_\alpha\boldsymbol{\beta} + \mathbf{Z})(\mathbf{y}_\alpha\boldsymbol{\beta} + \mathbf{Z})\lambda_\alpha - 2\boldsymbol{\lambda}_\alpha^T(\mathbf{y}_\alpha\boldsymbol{\beta} + \mathbf{Z})(\mathbf{y}(\mathbf{x})^T\boldsymbol{\beta} + Z(\mathbf{x}))... \\
&+ (\mathbf{y}(\mathbf{x})^T\boldsymbol{\beta} + Z(\mathbf{x}))^2] \\
=& (\boldsymbol{\lambda}_\alpha^T\mathbf{y}_\alpha\boldsymbol{\beta} - \mathbf{y}^T\boldsymbol{\beta})^2 + \sigma_z^2\boldsymbol{\lambda}_\alpha^T\mathbf{R}_\alpha\boldsymbol{\lambda}_\alpha + \sigma_z^2 - 2\sigma_z^2\boldsymbol{\lambda}_\alpha^T\mathbf{r}(\mathbf{x}) \\
=& \sigma_z^2(\boldsymbol{\lambda}_\alpha^T\mathbf{R}_\alpha\boldsymbol{\lambda}_\alpha + 1 - 2\boldsymbol{\lambda}_\alpha^T\mathbf{r}(\mathbf{x})) \\
=& \hat{\sigma}_g(\mathbf{x})
\end{aligned} \tag{3.13}
$$

其中$\mathbf{R}_\alpha$爲取樣與取樣之間的空間相關函數矩陣，其大小爲$n_s \times n_s$，$\mathbf{r}(\mathbf{x})$爲取樣與預測位置的空間相關函數矩陣，其大小爲$n_s \times 1$，分別如下

$$\mathbf{R}_\alpha = \begin{pmatrix} R(\mathbf{x}_1, \mathbf{x}_1) & R(\mathbf{x}_1, \mathbf{x}_2) & \cdots & R(\mathbf{x}_1, \mathbf{x}_{n_s}) \\ R(\mathbf{x}_2, \mathbf{x}_1) & & & \vdots \\ \vdots & & & \\ R(\mathbf{x}_{n_s}, \mathbf{x}_1) & & \cdots & R(\mathbf{x}_{n_s}, \mathbf{x}_{n_s}) \end{pmatrix}$$

$$\mathbf{r}(\mathbf{x}) = \begin{pmatrix} R(\mathbf{x}_1, \mathbf{x}) \\ R(\mathbf{x}_2, \mathbf{x}) \\ \vdots \\ R(\mathbf{x}_{n_s}, \mathbf{x}) \end{pmatrix}$$

至此我們已經擁有Kriging模型的推估變異數一般式，如方程式(3.13)，此一般式代表著Kriging模型與實際函數的方均誤差，因此我們要在無偏差拘束條件下，令其方均誤差最小化，因此可以將此問題寫成一般最佳化形式，如下方程式(3.14)。

$$\begin{aligned} \min_{\boldsymbol{\lambda}_\alpha} \quad & \sigma_z^2(\boldsymbol{\lambda}_\alpha^T \mathbf{R}_\alpha \boldsymbol{\lambda}_\alpha + 1 - 2\boldsymbol{\lambda}_\alpha^T \mathbf{r}(\mathbf{x})) \\ \text{s.t.} \quad & \mathbf{y}_\alpha^T \boldsymbol{\lambda}_\alpha = \mathbf{y}(\mathbf{x}) \end{aligned} \tag{3.14}$$

方程式(3.14)的最佳化問題可用拉格朗日方程式(Largrange Function)(3.15)來求解，其中ν爲拉格朗日因子(Largrange Multipliers)。

$$L_\nu = \sigma_z^2(\boldsymbol{\lambda}_\alpha^T \mathbf{R}_\alpha \boldsymbol{\lambda}_\alpha + 1 - 2\boldsymbol{\lambda}_\alpha^T \mathbf{r}(\mathbf{x})) - \nu[\mathbf{y}_\alpha^T \boldsymbol{\lambda}_\alpha - \mathbf{y}(\mathbf{x})] \tag{3.15}$$

將拉格朗日方程式對λ_α微分可得

$$\sigma_z^2 \mathbf{R}_\alpha \boldsymbol{\lambda}_\alpha - \sigma_z^2 \mathbf{r}(\mathbf{x}) - \mathbf{y}\nu = 0 \tag{3.16}$$

方程式(3.11)與方程式(3.16)之聯立解爲$\boldsymbol{\lambda}_\alpha$與$\nu$，將其寫成矩陣形式，如方程式(3.17)。

$$\begin{bmatrix} \sigma_Z^2\mathbf{R}_\alpha & \mathbf{y}_\alpha \\ \mathbf{y}_\alpha^T & \mathbf{0} \end{bmatrix} \begin{bmatrix} \boldsymbol{\lambda}_\alpha \\ -\nu \end{bmatrix} = \begin{bmatrix} \sigma_z^2\mathbf{r}(\mathbf{x}) \\ \mathbf{y}(\mathbf{x}) \end{bmatrix} \tag{3.17}$$

至此已可以利用方程式(3.17)的矩陣計算λ_α，而Kriging Estimator即可求得如方程式(3.18)，而Kriging模型的變異數即爲方程式(3.13)，將其改寫成矩陣形式，如方程式(3.19)。

$$\hat{g}(\mathbf{x}) = \boldsymbol{\lambda}_\alpha^T\mathbf{g}_\alpha = \begin{bmatrix} \mathbf{r}(\mathbf{x})^T & \mathbf{y}(\mathbf{x})^T \end{bmatrix} \begin{bmatrix} \mathbf{R}_\alpha & \mathbf{y}_\alpha \\ \mathbf{y}_\alpha^T & \mathbf{0} \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{g}_\alpha \\ \mathbf{0} \end{bmatrix} \tag{3.18}$$

$$\begin{aligned} \hat{\sigma}_g(\mathbf{x}) &= \sigma_z^2(1 + \boldsymbol{\lambda}_\alpha^T\mathbf{R}_\alpha\boldsymbol{\lambda}_\alpha - 2\boldsymbol{\lambda}_\alpha^T\mathbf{r}(\mathbf{x})) \\ &= \sigma_z^2\left[1 - \begin{bmatrix} \mathbf{r}(\mathbf{x})^T & \mathbf{y}(\mathbf{x})^T \end{bmatrix} \begin{bmatrix} \mathbf{R}_\alpha & \mathbf{y}_\alpha \\ \mathbf{y}_\alpha^T & \mathbf{0} \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{r}(\mathbf{x}) \\ \mathbf{y}(\mathbf{x}) \end{bmatrix}\right] \end{aligned} \tag{3.19}$$

根據文獻 [34]，可以將方程式(3.18)(3.19)改寫成另一種形式，如方程式(3.20)(3.21)。

$$\hat{g}(\mathbf{x}) = \mathbf{y}(\mathbf{x})^T\hat{\beta} + \mathbf{r}(\mathbf{x})^T\mathbf{R}_\alpha^{-1}(\mathbf{g}_\alpha - \mathbf{y}_\alpha\hat{\beta}) \tag{3.20}$$

$$\hat{\beta} = (\mathbf{y}_\alpha^T\mathbf{R}_\alpha^{-1}\mathbf{y}_\alpha)^{-1}\mathbf{y}_\alpha^T\mathbf{R}_\alpha^{-1}\mathbf{g}_\alpha$$

$$\hat{\sigma}_g(\mathbf{x}) = \sigma_z^2(1 - \mathbf{r}(\mathbf{x})^T\mathbf{R}_\alpha^{-1}\mathbf{r}(\mathbf{x})) \tag{3.21}$$

以上所敘述的四個方程式(3.18-3.21)，均可得到Kriging模型的平均值與變異數，差別只在於後者多計算了母體估計平均值$\hat{\beta}$。而在此處我們所得到的Kriging模型，稱之爲Universal Kriging，如果上述的方程式中的回歸函數$y_j(\mathbf{x})$等於1，則我們稱之爲Ordinary Kriging，如方程式(3.22)(3.23)，Ordinary Kriging具有一個特色，即利用無偏差拘束條件可以看出權重的總和等於一，如方程式(3.24)。

$$
\begin{aligned}
\hat{g}(\mathbf{x}) &= \begin{bmatrix} \mathbf{r}(\mathbf{x})^T & \mathbf{1} \end{bmatrix} \begin{bmatrix} \mathbf{R}_\alpha & \mathbf{1} \\ \mathbf{1} & \mathbf{0} \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{g}_\alpha \\ \mathbf{0} \end{bmatrix} \\
&= \mathbf{1}\hat{\beta} + \mathbf{r}(\mathbf{x})^T \mathbf{R}_\alpha^{-1}(\mathbf{g}_\alpha - \mathbf{1}\hat{\beta})
\end{aligned} \tag{3.22}
$$

$$
\hat{\beta} = (\mathbf{1}\mathbf{R}_\alpha^{-1}\mathbf{1})^{-1}\mathbf{1}\mathbf{R}_\alpha^{-1}\mathbf{g}_\alpha
$$

$$
\begin{aligned}
\hat{\sigma}_g(\mathbf{x}) &= \sigma_z^2 \left[1 - \begin{bmatrix} \mathbf{r}(\mathbf{x})^T & \mathbf{1} \end{bmatrix} \begin{bmatrix} \mathbf{R}_\alpha & \mathbf{1} \\ \mathbf{1} & \mathbf{0} \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{r}(\mathbf{x}) \\ \mathbf{1} \end{bmatrix} \right] \\
&= \sigma_z^2 (1 - \mathbf{r}(\mathbf{x})^T \mathbf{R}_\alpha^{-1} \mathbf{r}(\mathbf{x}))
\end{aligned} \tag{3.23}
$$

$$
\mathbf{1}^T \boldsymbol{\lambda}_\alpha(\mathbf{x}) = 1 \tag{3.24}
$$

Ordinary Kriging是目前工程領域最常使用的Kriging模型，本論文程式內容將以Ordinary Kriging爲主。然而以上方程式的堆導，都是在函數的共變異數方程式(3.12)是在已知的情況下所得到的結果，然而不同的共變異數方程式將會產生截然不同的Kriging模型，關於如何決定此共變異數方程式將在下一節說明。

3.3 參數選擇 (Parameter Estimation)

如同上一節的說明，不同的共變異數方程式將會產生截然不同的Kriging模型，因此決定此方程式是一件重要的事，而我們稱這個過程爲擬合(Fitting)，現今對Kriging模型進行擬合的方法主要有三種，分別是變異圖分析(Variogram)、最大似然推估法(Maximal Likelihood Estimation)以及交叉驗證法(Cross Validation)。

變異圖分析源自於地理統計學，是最初原本的擬合方法；最大似然推估法被廣泛應用於工程領域，包含了EGO演算法；而交叉驗證法的概念相當簡單，與單點移出法(Leave One Out)相似，先將所有取樣中抽出少數取樣，定義一個共變異數方程式並且設定此共變異數方程式的參數，然後將沒有被抽出的取樣與此方程式建立Kriging模型，藉由不同方程式參數來比較Kriging模型與抽出的取樣之間的誤差，之後再重複的取出不同的取樣並執行上述動作，在文獻 [35]中有比較最大似然推估法與交叉驗證法，然而此方法的擬合時間較長，因此比較少有文獻將其作爲演算法的擬合方法，大部分使用交叉驗證法是爲了要判斷目前Kriging模型與原函數誤差多少，因此在本論文中，將以變異圖分析與最大似然推估法作爲討論的主題。

3.3.1 變異圖分析 (Variogram)

變異圖是一種統計上的量測工具，用來量測一個隨機函數的空間相關程度，其詳細內容可參考文獻 [32,36]，其數學定義與本質假設的第二個方程式(3.7)相同，如下列方程式(3.25)，其中$2\gamma(h)$爲變異圖。

$$2\gamma(h) = E[(G(\mathbf{x}_i) - G(\mathbf{x}_j))^2] \tag{3.25}$$

考慮一個滿足平穩假設與本質假設的隨機函數$G(\mathbf{x})$，則變異圖方程式可改寫成方程式(3.26)，其中$\gamma(h)$稱之爲半變異圖(Semivariogram)

$$
\begin{aligned}
2\gamma(h) = E[(G(\mathbf{x}_i) - G(\mathbf{x}_j))^2] &= E[G(\mathbf{x}_i)^2 - 2G(\mathbf{x}_i)G(\mathbf{x}_j) + G(\mathbf{x}_j)^2] \\
\because E[G(\mathbf{x}_i)^2] &= E[G(\mathbf{x}_j)^2] = \sigma_z^2 \\
E[G(\mathbf{x}_i)G(\mathbf{x}_j)] &= Cov(G(\mathbf{x}_i), G(\mathbf{x}_j)) = Cov(Z(\mathbf{x}_i), Z(\mathbf{x}_j)) \\
\therefore E[G(\mathbf{x}_i)^2 - 2G(\mathbf{x}_i)G(\mathbf{x}_j) + G(\mathbf{x}_j)^2] &= 2\sigma_z^2 - 2Cov(Z(\mathbf{x}_i), Z(\mathbf{x}_j)) \\
\therefore \gamma(h) &= \sigma_z^2 - Cov(Z(\mathbf{x}_i), Z(\mathbf{x}_j)) \qquad (3.26)
\end{aligned}
$$

至此，我們已經擁有變異圖與共變異數方程式之間的關係，之所以推導變異圖與共變異數之間的關係，其最大用意是一個隨機歷程的共變異數是無法直接由取樣得知，而變異圖卻可以經由取樣所得到的變異圖的計算以觀察其趨勢。

因此如何由取樣來計算變異圖呢？由於實際量測的數據皆爲定率資料，所以我們將變異圖方程式(3.25)改寫成方程式(3.27)，我們稱此式爲經驗變異圖(Empirical Variogram)。

$$
2\gamma(h) = \frac{1}{N(h)} \sum_{N(h)} [g(\mathbf{x}) - g(\mathbf{x} + h)]^2 \qquad (3.27)
$$

其中$N(h)$代表取樣之間距離爲h的樣本個數，由經驗變異圖方程式的內容來說明，即爲在所有的樣本中，符合樣本之間的距離爲h的樣本，將其量測值兩兩相減平方總和之後，再除以距離爲h的樣本總數。可參考圖3.3，藍色x爲樣本，距離h_1與h_2爲不同距離的樣本，所以必須分開計算，而滿足距離爲h_1的樣本有$N(h_1)$個，滿足距離爲h_2的樣本有$N(h_2)$個，以此類推，以下將以一個範例來進行說明。

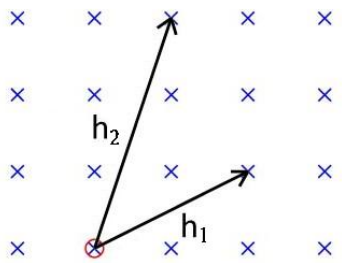
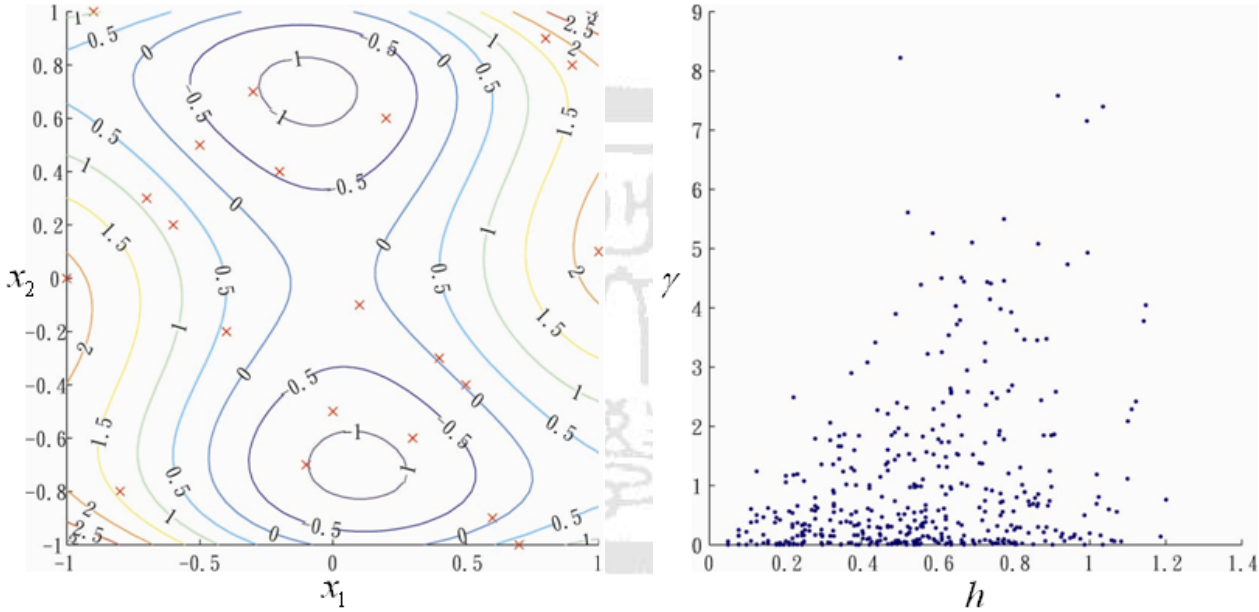


圖 3.3: 取樣與距離示意圖

- **範例：**Camelback**函數** [2]

$$f(\mathbf{x}) = -0.1 + \left(4 - 2.1x_1^2 + \frac{x_1^4}{3}\right) x_1^2 + x_1 x_2 + \left(-4 + 4x_2^2\right) x_2^2 \tag{3.28}$$

此範例爲一個二維的函數，如圖3.4**左圖，其中等高線即爲函數值，**x**即爲取樣位置。根據此函數與取樣，依照經驗變異圖方程式**(3.27)**計算後可以得到圖**3.4**右圖，我們稱這張圖爲變異圖雲圖**(Variogram Cloud)**。**



(a) Camelback**函數取樣圖**　　(b) Camelback**變異圖雲圖**

圖 3.4: **靠近原點之兩個說明範例樣本圖**

然而此變異雲圖資料過於分散，仍然無法看出變異圖的趨勢，因此我們使用實驗變異圖(Experimental Variogram)，其概念是將經驗變異圖方程式的距離h，改爲滿足一個區間範圍內的樣本，如圖3.5，將其量測值兩兩相減平方總和之後，再除以滿足區間範圍內的樣本總數，然後將此點實驗變異圖資訊的距離定爲此區間的一半，也可以定爲區間內所有樣本的平均距離，如方程式(3.29)，將此範例之實驗變異圖繪製成圖，如圖3.6。

$$2\gamma(h) = \frac{1}{N(h_k)} \sum_{N(h_k)} [g(\mathbf{x}_i) - g(\mathbf{x}_j)]^2 \tag{3.29}$$

$$h_k^l \leq \|\mathbf{x}_i - \mathbf{x}_j\| \leq h_k^u$$

$$h_k = \frac{1}{2}(h_k^l + h_k^u)$$

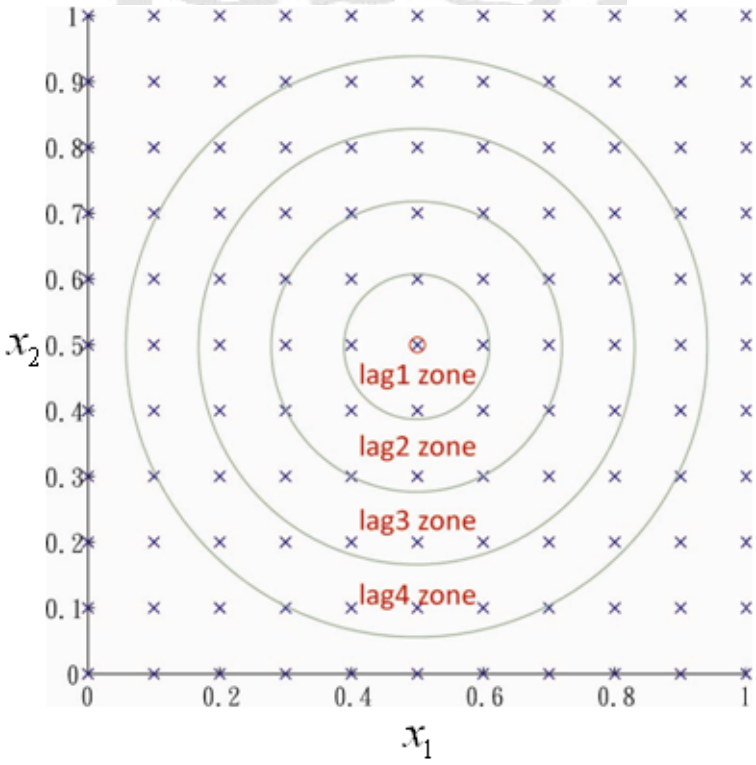


圖 3.5: 距離區域範圍示意圖

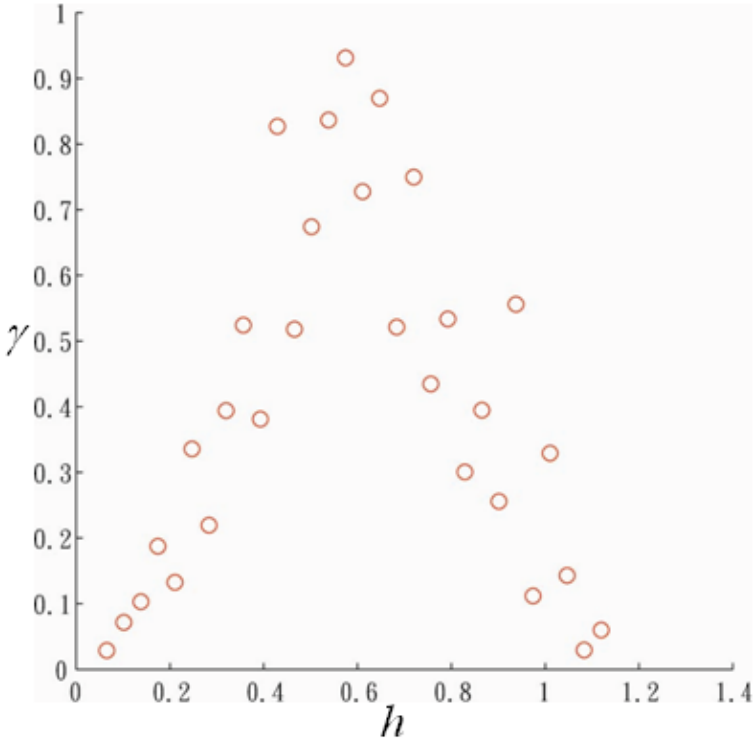


圖 3.6: Camelback實驗變異圖圖

至此便完成了實驗變異圖的計算，然而藉由此實驗變異圖可以得到此函數的什麼資訊呢？在本小節一開始提到過，變異圖是一種統計上的量測工具，用來量測一個隨機函數的空間相關程度，因此實驗變異圖在圖形上有其數值的統計意義，其主要可以分成兩方面，分別是微觀現象以及大趨勢。

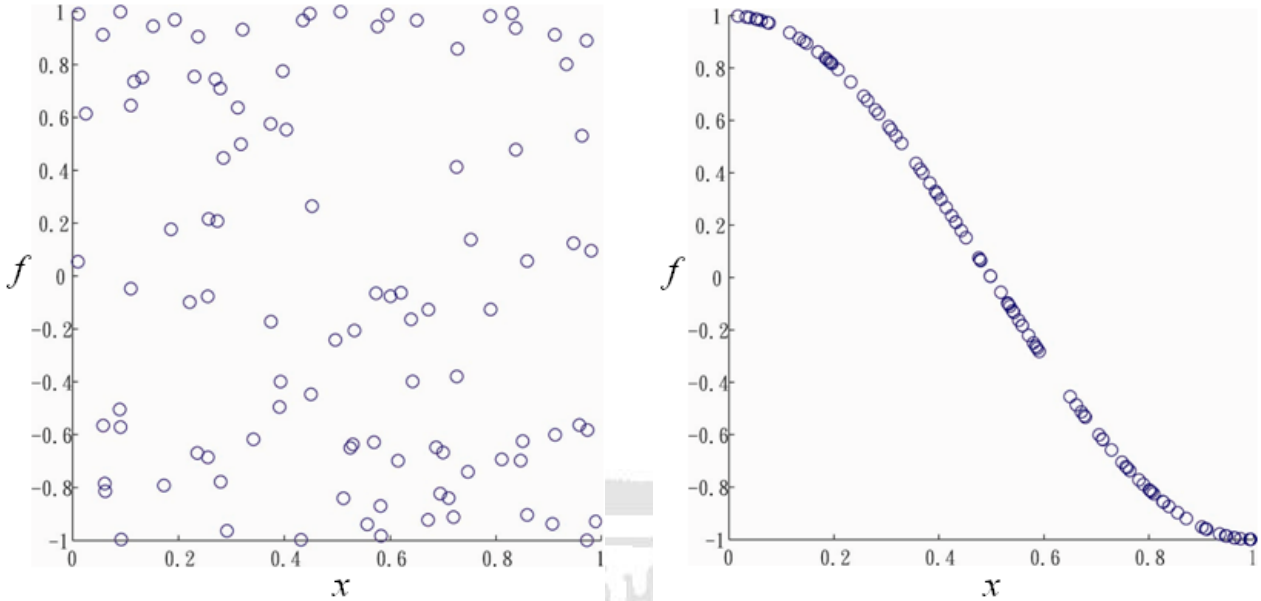
- 微觀現象

以下用兩個範例來說明實驗變異圖圖在靠近原點的行爲，分別是

$$f(x) = \cos\left(\frac{2\pi x}{0.001}\right)$$

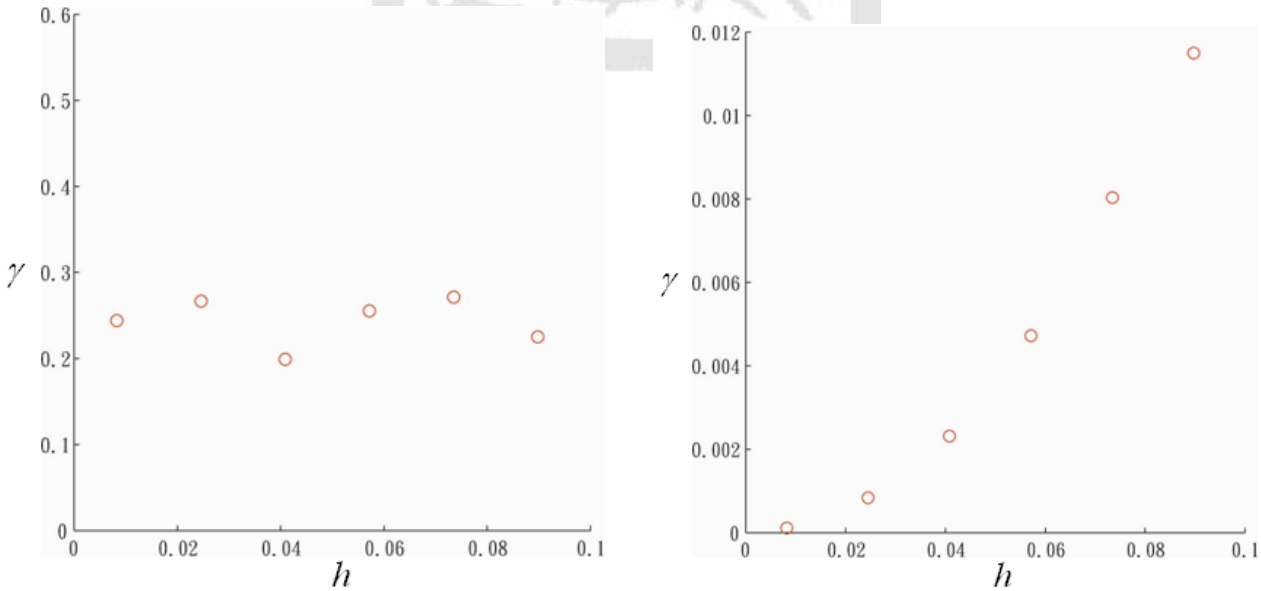
$$f(x) = \cos(\pi x)$$

分別對此兩個函數隨機取樣100個不重複樣本，將其繪製如圖3.7。由樣本圖可以發現，$f(x) = \cos\left(\frac{2\pi x}{0.001}\right)$在取樣範圍內是一個持續震盪、且似乎不具有連續性質的函數，而$f(x) = \cos(\pi x)$是一個平滑的曲線，將此兩種函數的樣本經過實驗變異圖計算後，將實驗變異圖靠近原點處繪製成圖3.8。



(a) $f(x) = \cos\left(\frac{2\pi x}{0.001}\right)$ (b) $f(x) = \cos(\pi x)$

圖 3.7: 靠近原點之兩個說明範例樣本圖



(a) $f(x) = \cos\left(\frac{2\pi x}{0.001}\right)$ (b) $f(x) = \cos(\pi x)$

圖 3.8: 靠近原點處之兩個說明範例實驗變異圖

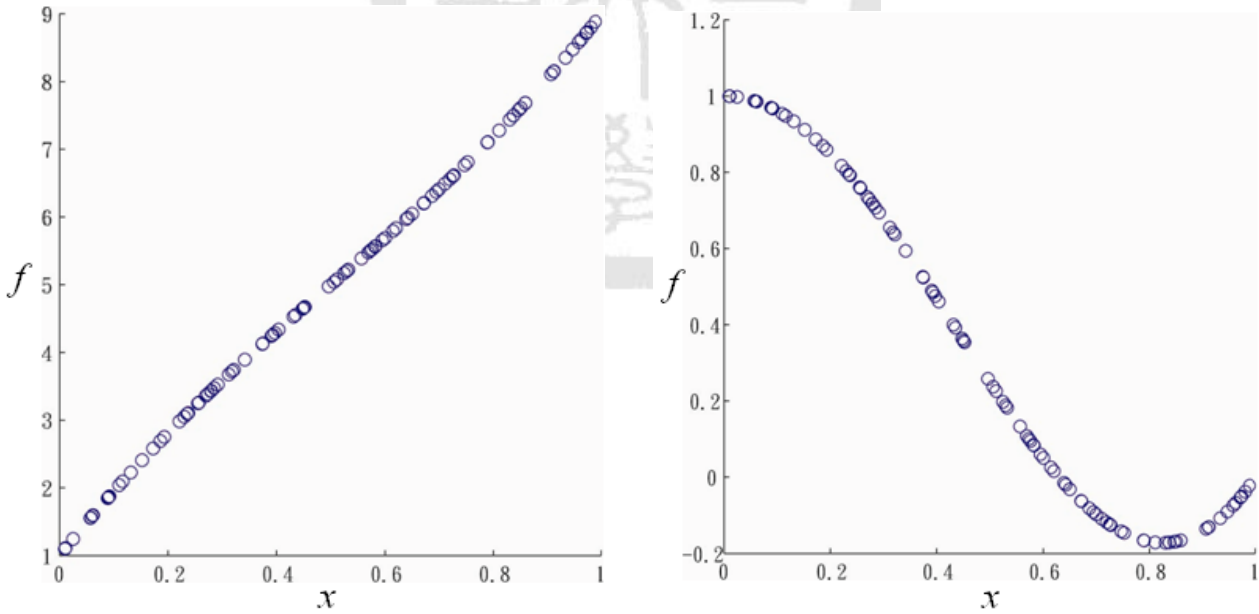
由兩者的實驗變異圖靠近原點處的圖形可以發現，函數$f(x) = \cos\left(\frac{2\pi x}{0.001}\right)$的實驗變異圖之延伸並不會通過原點，而函數$f(x) = \cos(\pi x)$的實驗變異圖之延伸不僅通過原點，並且呈現了拋物線狀態，因此我們可以下一個結論，即爲實驗變異圖在靠近原點處的行爲，即代表著此函數在取樣的範圍內的連續性與平滑程度。

- 巨觀趨勢

以下用兩個範例來說明實驗變異圖在巨觀趨勢的行爲，分別是

$$f(x) = 10x + \cos(\pi x)$$
$$f(x) = x^2 + \cos(\pi x)$$

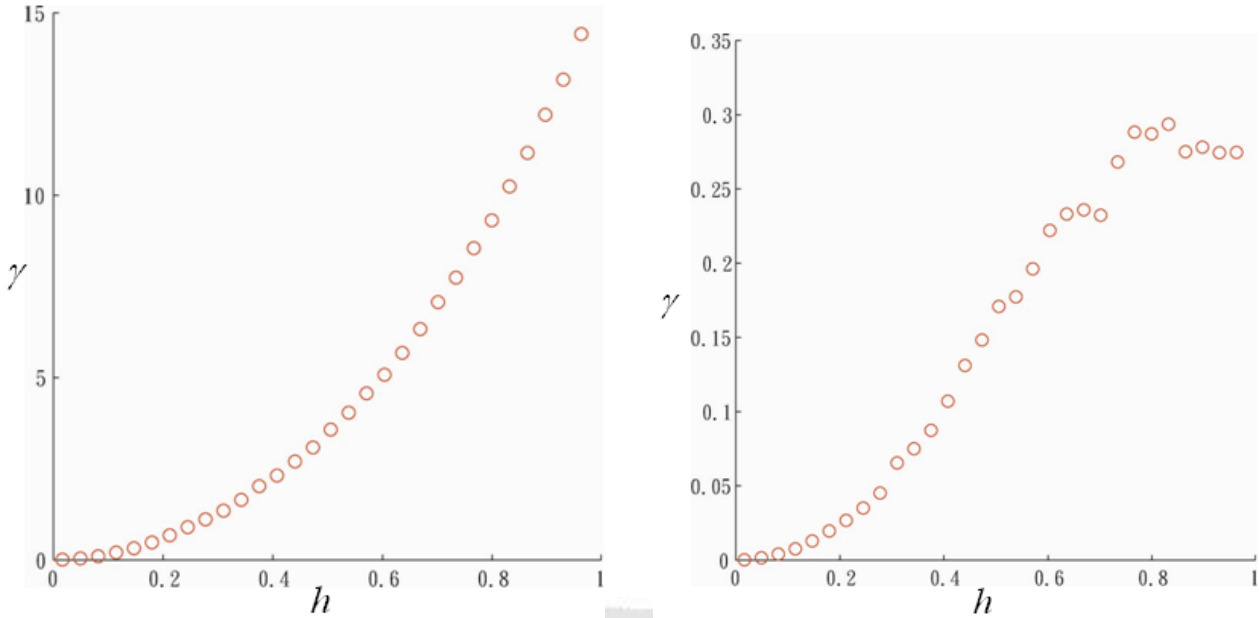
分別對此兩個函數隨機取樣100個不重複樣本，將其繪製如圖3.9。



(a) $f(x) = 10x + \cos(\pi x)$ (b) $f(x) = x^2 + \cos(\pi x)$

圖 3.9: 巨觀趨勢之兩個說明範例樣本圖

由樣本圖可以發現，$f(x) = 10x + \cos(\pi x)$與$f(x) = x^2 + \cos(\pi x)$在取樣範圍內皆爲一個平滑的曲線，然而兩者最大的差異是，$f(x) = 10x + \cos(\pi x)$在取樣空間內是一個單調性函數(Monotonic Function)，將此兩種函數的樣本經過實驗變異圖計算後，將實驗變異圖的大趨勢，分別繪製成圖3.10。



(a) $f(x) = 10x + \cos(\pi x)$ (b) $f(x) = x^2 + \cos(\pi x)$

圖 3.10: 巨觀趨勢之兩個說明範例實驗變異圖

由兩者的實驗變異圖圖形的巨觀趨勢可以發現，函數$f(x) = 10x + \cos(\pi x)$的實驗變異圖在距離越大的情況下持續的增加，而函數$f(x) = x^2 + \cos(\pi x)$的實驗變異圖在某個距離之後不再持續上升，我們將此兩種大趨勢的行爲成爲平穩(Stationary)與非平穩(Nonstationary)，如圖3.11。

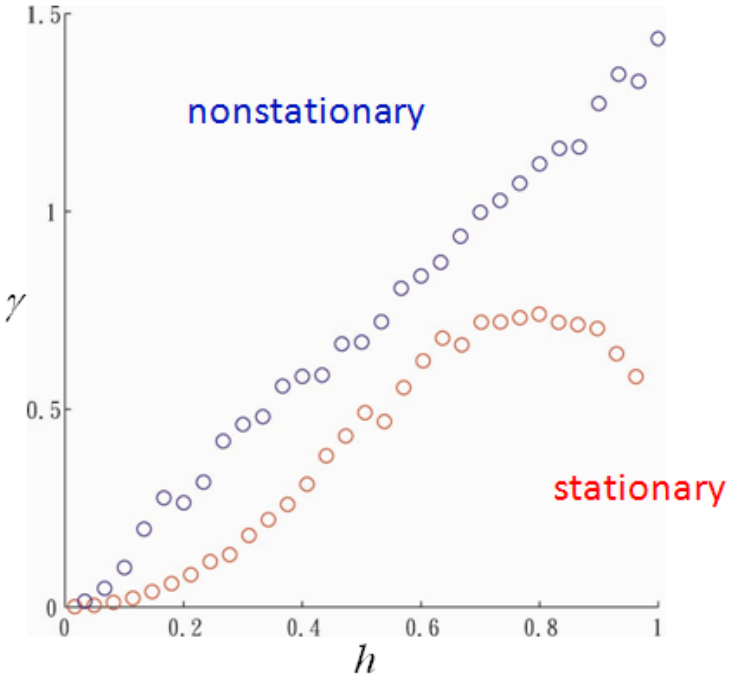


圖 3.11: 平穩與非平穩實驗變異圖示意圖

由上圖可知，我們將實驗變異圖在某個距離之後不再持續上升的這個行爲定義爲平穩，在統計上的意義是此隨機函數在超過某一個距離之後，其不同位置的相關性會變得很小，也可以說是不相關(Uncorrelated)，而非平穩是無論隨機函數的距離爲多少皆爲相關，大部分的自然現象應屬於平穩狀態，然而會出現非平穩的情形也並不代表此自然現象爲非平穩，很有可能是在取樣的空間範圍內所收集到的資訊爲非平穩。

因此在一個平穩的實驗變異圖中，當實驗變異圖到達穩定狀態下圖形高度，稱之爲實驗變異圖的尾檻值(Sill)，而到達這山脊的距離稱之爲實驗變異圖的影響範圍(Influence Range)或是相關距離(Correlation Length)，另外在接近原點處的高度稱之爲碎塊效應(Nugget Affect)，如圖3.12。

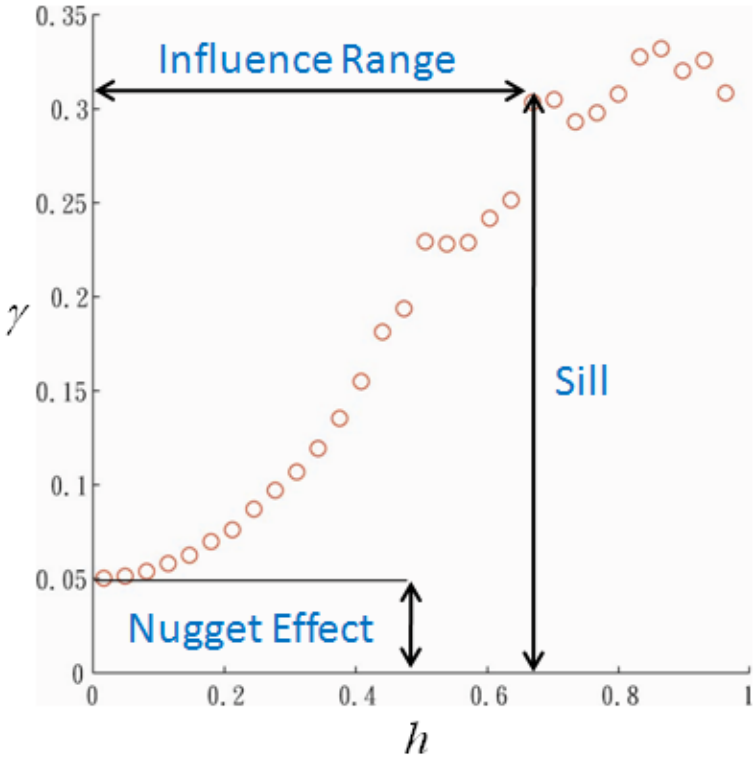


圖 3.12: 平穩型實驗變異圖名詞示意圖

至此，我們大致上已經知道究竟如何利用變異圖來瞭解一個隨機函數在空間中不同位置的相關性，然而由實驗變異圖所得到的資訊是離散的，但是我們必須對Kriging模型上的任何位置皆有定義其相關性，因此需要制定一些數學式來取代這些由實驗變異圖所得到的資訊，而我們稱這些數學式爲理論變異圖(Theoretical Variogram)，以下將列出三種常用的理論變異圖。

- 球形模式(Spherical Model)

$$\gamma(h) = \begin{cases} C_0 + \sigma_z^2 \left(\frac{3h}{2a} - \frac{h^3}{2a^3} \right), & 0 \le h \le a \\ C_0 + \sigma_z^2, & h > a \\ 0, & h = 0 \end{cases} \tag{3.30}$$

$$R(h) = \begin{cases} 1 - \frac{3h}{2a} - \frac{h^3}{2a^3}, & 0 \le h \le a \\ 0, & h > a \end{cases} \tag{3.31}$$

將方程式中的參數類比爲實驗變異圖中的意義，則C_0代表著碎塊效應，σ_z^2代表整體

變異數，$C_0+\sigma_z^2$即爲山脊，a代表影響範圍，$R(h)$爲相對應的空間相關函數，而這些參數即爲Kriging模型中需要進行擬合的參數，即爲本小節主要的討論內容。

- **指數模式**(Exponential Model)

$$\gamma(h)=\begin{cases} C_0+\sigma_z^2\left(1-\exp\left(-\frac{h}{L}\right)\right), & h>0 \\ 0, & h=0 \end{cases} \tag{3.32}$$

$$R(h)=\begin{cases} \exp\left(-\frac{h}{L}\right), & h>0 \\ 0, & h=0 \end{cases} \tag{3.33}$$

方程式中的影響範圍大約等於$3L$，其他與球形模式敘述無異。

- **高斯模式**(Gaussian Model)

$$\gamma(h)=\begin{cases} C_0+\sigma_z^2\left(1-\exp\left(-\frac{h^2}{L^2}\right)\right), & h>0 \\ 0, & h=0 \end{cases} \tag{3.34}$$

$$R(h)=\begin{cases} \exp\left(-\frac{h^2}{L^2}\right), & h>0 \\ 0, & h=0 \end{cases} \tag{3.35}$$

方程式中的影響範圍大約等於$\frac{7L}{4}$，其他與球形模式敘述無異。

在地質統計領域中，因爲希望能保留使用者對於Kriging模型調整的彈性，因此對於Kriging模型的擬合並沒有一個絕對標準的作法，一般來說都是將實驗變異圖繪製圖形後，觀察變異圖的行爲與趨勢，並結合此自然現象的特性來設定理論變異圖的參數，以達到建立一個更適合此自然現象的推估模型的目的。

然而對於EGO演算法來說，我們需要一個能自動化的流程來達到演算的目的，因此在本論文的程式中關於變異圖擬合的部份，是利用實驗變異圖與理論變異圖之間的平方誤差最小化的概念來進行擬合，如方程式(3.36)，其中γ_α代表實驗變異圖、γ_t代表著理論變異圖，而計算後所選用來擬合的實驗變異圖有k個，而每個實驗變異圖皆有其對應之距離。

$$\min_{C_0,\theta,\sigma_z^2,p} \quad \sqrt{\sum_{i=1}^{k}\left[\gamma_\alpha(h_i)-\gamma_t(h_i)\right]^2} \tag{3.36}$$

必須注意到我們在使用實驗變異圖來進行擬合時，並不是將所有的實驗變異圖都視爲有用的資訊，因爲當距離變大時，能夠蒐集到的變異圖資訊的點數將變少，很有可能會造成實驗變異圖的偏差，因此在本論文中，我們使用所有實驗變異圖中最大距離的一半，來作爲擬合時選取實驗變異圖資訊的範圍。

爲了與最大似然推估法(Maximal Likelihood Estimation)在程式中達成一致性，此方法將在下一小節作介紹，我們所選用的理論變異圖是一種類似於韋氏分佈概念的數學式，如方程式(3.37)，其相對應的空間相關函數$R(h)$，如方程式(3.38)，其中C_0,θ,σ_z^2,p皆爲擬合之參數，其本身也是方程式(3.36)最佳化問題之變數。

$$\gamma(h)=\begin{cases} C_0+\sigma_z^2\left(1-\exp(\theta h^p)\right), & h>0 \\ 0, & h=0 \end{cases} \tag{3.37}$$

$$R(h)=\exp(\theta h^p) \tag{3.38}$$

然而並非所有參數皆需擬合，使用者可以自行決定Kriging模型所要表現的型態，例如若是不想重現碎塊效應的性質，則可直接將C_0定爲常數，可參考演算程式 *superEGO* 的文獻 [33]，演算程式 *superEGO* 是由美國密西根大學Sasena博士所撰寫的一種EGO演算法程式，之所以會稱之爲super的原因是因爲此程式是針對黑盒子函數與純數學函數的各種擴展型問題(Superset)所撰寫，Sasena爲了改善DACE所使用的Kriging模型，將方程式(3.37)中的p之預設值定爲1.99，C_0定爲10^{-12}，這使得方程式中需要擬合的變數剩下θ與σ_z^2。至此我們已完成了Kriging模型的自動擬合方法說明，而我們稱此方法爲等向性最小變異圖平方誤差擬合(Least Square Isotropic Variogram Fitting)，簡稱爲最小變異圖誤差。

同樣以Camelback函數 [2]作爲範例，如方程式(3.28)，取樣位置與圖3.4相同，擬合完成後將實驗變異圖與理論變異圖結果繪製，如圖3.13，而原函數與Kriging模型重建預測結果比較如圖3.14。

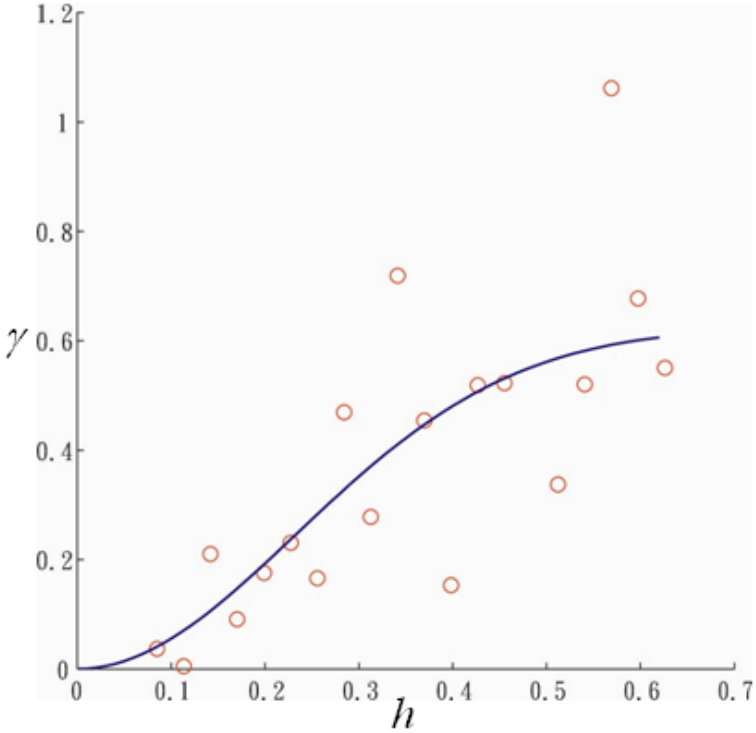
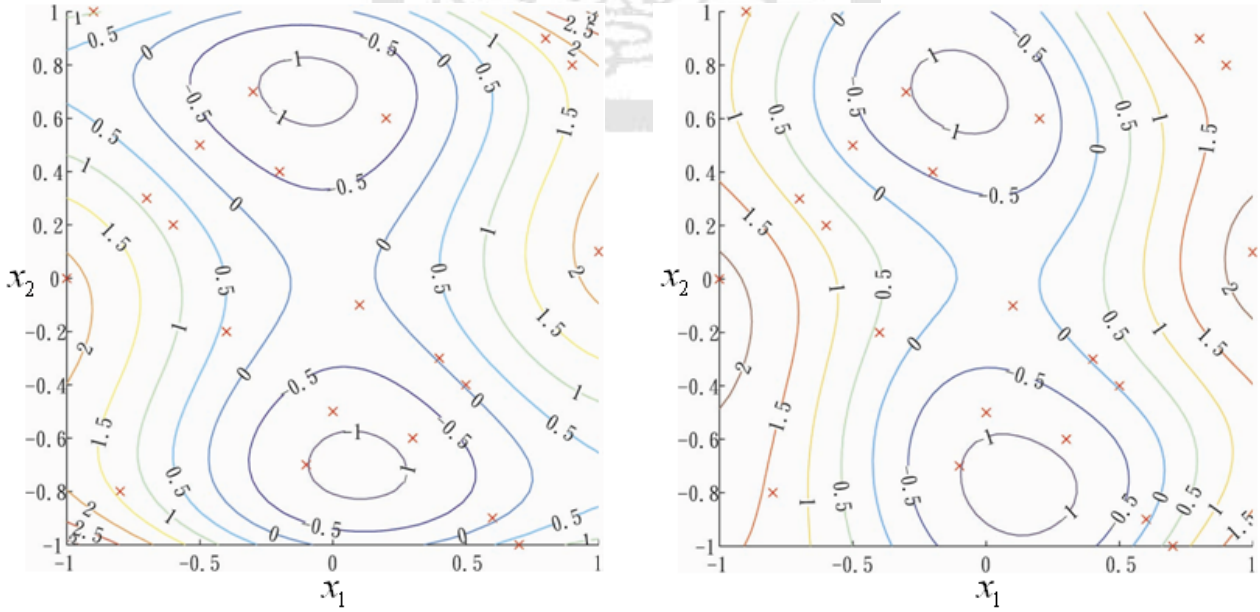


圖 3.13: 實驗變異圖與理論變異圖之擬合



(a) 原函數與取樣位置　　(b) 最小變異圖誤差擬合結果

圖 3.14: 最小變異圖誤差擬合結果函數比較圖

本小節只將變異圖概念與內容做一個簡單的介紹，事實上在使用變異圖進行擬合

時，還可以考慮其他統計上的性質，如等向性(Isotropic)、非等向性(Anisotropic)、週期性(Cyclicity)等，欲瞭解有關於更多變異圖的概念，可參考 [36]。

3.3.2 最大似然推估法 (Maximal Likelihood Estimation)

最大似然推估法是由著名的英國科學家Fisher於1920年代所提出，是一種用來擬合資料與某種統計模型參數的數學方法，此方法在各種統計相關書籍皆有介紹 [7,37]。

在1989年之後，最大似然推估法在地質統計學以外的領域得到了廣泛的應用，相當多的範例在使用Kriging模型時，都會使用最大似然推估法來作爲擬合的手段 [27–30]，而這些文獻中大多都將此方法引用自DACE [25,26,38]，然而眞正最早開始針對一個高斯歷程使用似然函數(Likelihood Function)求取參數的文獻，可追溯到更早的文獻 [39,40]。而在1998年，由Jones、Schonlau、Welch等人所發展的EGO演算法 [10]，以及後來研究EGO演算法的文獻 [41–45]，也大多使用最大似然推估法來進行擬合。

對一個高斯分佈的隨機函數X來說，欲求得此隨機函數的平均值與標準差，如果有n個數據$x_1, x_2, ...x_n$是由此隨機函數所取樣產生，取樣之間爲獨立(Independent)關係，假設此隨機函數的機率密度函數(Probability Density Function)爲$f_X(x, \mu, \sigma)$，則似然函數可定義爲方程式(3.39)。

$$L(x_1, x_2, ..., x_n, \mu, \sigma) = f_X(x_1, \mu, \sigma) f_X(x_2, \mu, \sigma) .... f_X(x_n, \mu, \sigma) \tag{3.39}$$

如果此隨機函數不只限於高斯分佈，其機率密度函數爲$f_X(x, p)$，其中p爲此隨機函數之參數，則方程式(3.39)可改寫成方程式(3.40)。

$$L(x_1, x_2, ..., x_n, p) = f_X(x_1, p) f_X(x_2, p) .... f_X(x_n, p) \tag{3.40}$$

根據此似然函數的參數，對此函數進行最大化，我們稱這個過成爲最大似然推估法(Maximum Likelihood Estimation)，可以直接利用似然函數的一次微分求得，如$\frac{\partial L}{\partial p} = 0$，因此，簡單的說最大似然推估法就是根據目前的樣本，在設定某種已知的分佈情形下所猜測的參數，而此組參數是此分佈最有可能重現目前樣本的參數，因此才稱之爲最大似然。而在DACE中，對於Kriging模型所使用的空間相關函數與方程式(3.38)類似，與之前不一樣的地方是DACE將各個維度的變數分開計算，形成一種具有非等向性(Anisotropy)空間相

關函數的效果，然後使用 *product correlation rule* 結合爲一個空間相關函數，此規則可參考 [38]；根據文獻 [34]，我們將空間相關函數寫成方程式(3.41)，其中D代表著問題的維度總數。

$$R(\mathbf{x}_i, \mathbf{x}_j) = \prod_{d=1}^{D} \exp^{-\theta_d |x_{i_d} - x_{j_d}|^{p_d}} \tag{3.41}$$

根據文獻 [38,46]，利用貝氏方法(Bayesian Approach)在不考慮碎塊效應的情況下所得到的log-likelihood方程式，如方程式(3.42)。其中σ_z^2可直接利用$\frac{\partial L}{\partial \sigma_z^2} = 0$求得，如方程式(3.43)。

$$L(\sigma_z^2, \boldsymbol{\theta}, \mathbf{p}) = -\frac{1}{2}\left[n\ln(2\pi) + n\ln\sigma_z^2 + \ln\det(\mathbf{R}_\alpha) + \frac{1}{\sigma_z^2}(\mathbf{g}_\alpha - \mathbf{y}_\alpha\hat{\beta})^T\mathbf{R}_\alpha^{-1}(\mathbf{g}_\alpha - \mathbf{y}_\alpha\hat{\beta})\right] \tag{3.42}$$

$$\sigma_z^2 = \frac{1}{n}(\mathbf{g}_\alpha - \mathbf{y}_\alpha\hat{\beta})^T\mathbf{R}_\alpha^{-1}(\mathbf{g}_\alpha - \mathbf{y}_\alpha\hat{\beta}) \tag{3.43}$$

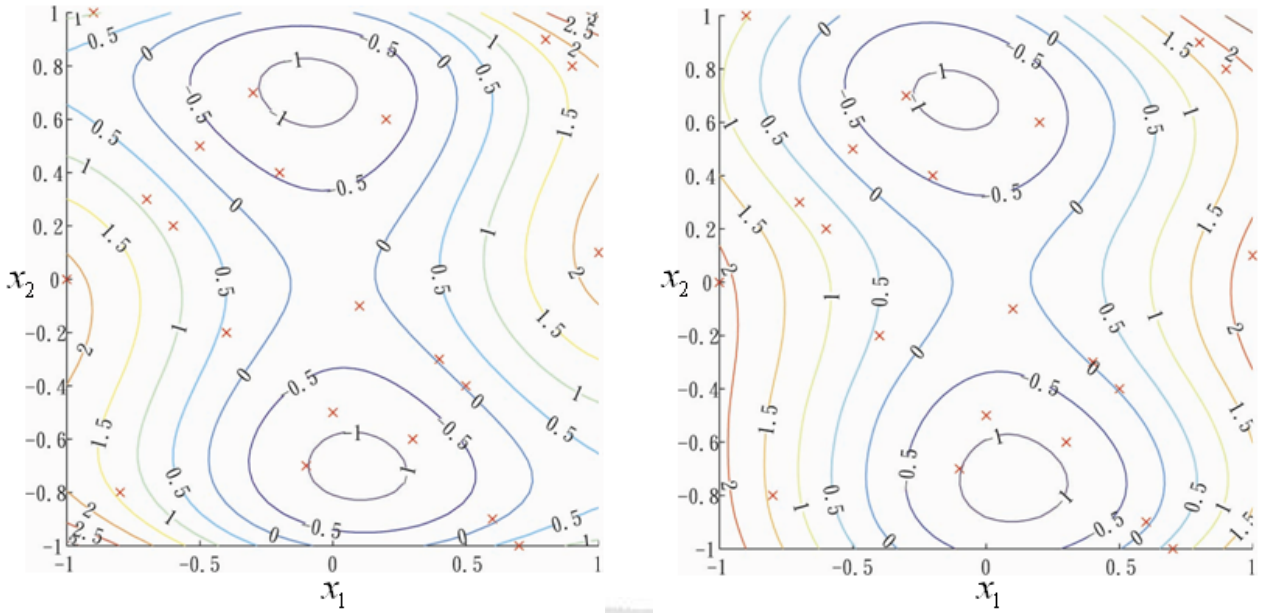
然而我們卻無法將似然函數直接對$\boldsymbol{\theta}$與$\mathbf{p}$微分來求取參數，因此我們必須依賴一些最佳化演算法，來進行此最大化似然函數的動作，而最佳化的目標函數，根據文獻 [34]，我們將上列方程式(3.43)改寫成方程式(3.44)，其中$\boldsymbol{\theta}, \mathbf{p}$是$\mathbf{R}_\alpha$的變數，至此我們已經可以利用最大可能性方法來對Kriging模型進行擬合，以下舉一範例說明。

$$L(\boldsymbol{\theta}, \mathbf{p}) = -\frac{1}{2}\left[n\ln\sigma_z^2 + \ln\det(\mathbf{R}_\alpha)\right] \tag{3.44}$$

同樣以Camelback函數 [2]作爲範例，參照方程式(3.28)，原函數及取樣位置與圖3.14相同，擬合結果如圖3.15，可以與最小變異圖誤差擬合的結果圖3.14作比較，不過在此並不進行兩種擬合方法之結果的比較，但是我們可以經由上述清楚瞭解到最大似然推估法與變異圖分析之不同，在此節錄一段自地理統計文獻 [36]中評論最大似然推估法的文章如下：

> *"Some methods work directly from sample data, without requiring the calculation of the sample variogram. Since these methods are blind, they tend to be used only when the presence of a strong drift causes the sample variogram to be hopelessly biased."*

其意思是在進行空間函數推估時，使用最大似然推估法而不計算變異圖，必須建立在所有的數據都是在變異圖極少偏差(Hopelessly Biased)的情況下才能使用，這是因爲最大似然推估法是將所有的數據都加入擬合的考慮。



(a) 原函數與取樣位置　　(b) 最大似然推估法擬合結果

圖 3.15: 最大似然推估法擬合結果函數比較圖

3.4 問題與討論 (Problem and Discussion)

在使用最大似然推估法進行參數擬合時，我們觀察到一個現象，就是當資料的樣本數增加時，演算法所找到的似然函數值會等於無限大，以下用一個一維範例來說明此現象。

- 範例：一維測試函數

$$f(x) = -\sin x - e^{\frac{x}{100}} + 10 \tag{3.45}$$

圖3.16爲其函數圖形與對此函數取網格化樣本後之示意圖。

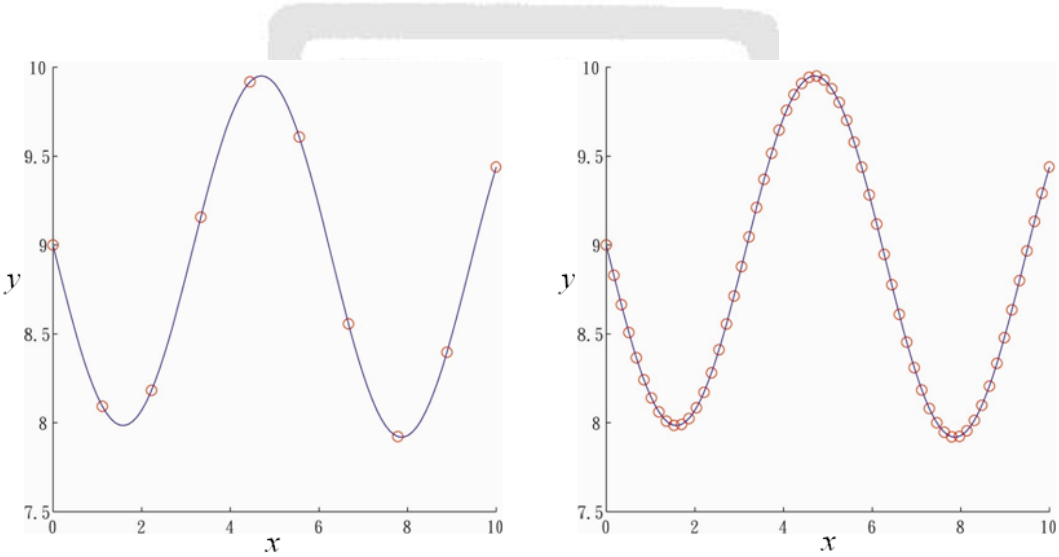


圖 3.16: 一維測試函數樣本示意圖

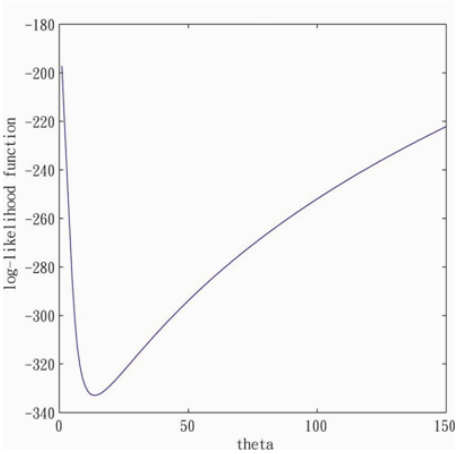
根據似然函數方程式(3.44)，並在空間相關係數使用方程式(3.41)，只考慮一個維度θ的情況下，不同的樣本數目執行似然函數最佳化的結果如表3.1。由表中最後兩次最佳化結果可以發現其目標函數數值皆爲無限大，然而此現象對於演算法來說並不正常，這代表目標函數在此時並不是一個有定義的數值，其原因發生在似然方程式(3.44)中關於$\ln\det(\mathbf{R}_\alpha)$的部份，當樣本數增加時，樣本的空間相關函數矩陣$\mathbf{R}_\alpha$的行列式值將越趨近於零，而零的自然對數爲負無限大。

表 3.1: 似然函數最佳化的結果

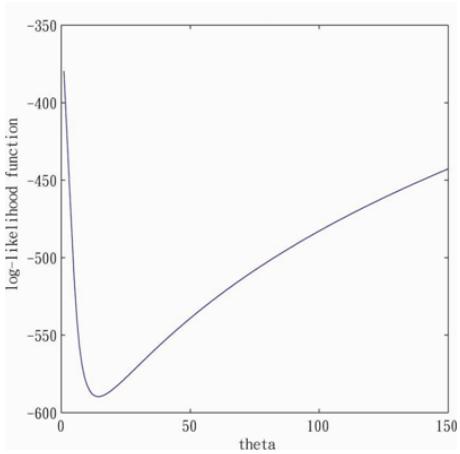
樣本數目	θ	log-likelihood函數值
10	11.0490	20.8389
15	11.6521	61.7892
20	12.2669	109.0999
25	12.7440	160.7000
30	13.1187	215.5844
35	13.4199	273.1412
40	13.6673	332.9549
45	13.8739	394.7236
50	14.0492	458.2176
55	0.0183	infinite
60	0.0610	infinite

我們在不同的樣本數下，將似然函數(Log-likelihood Function)與θ繪製，如圖3.17，注意由於演算法皆以最小化爲標準動作，因此此圖中的似然函數與方程式(3.44)相差負號，所以最佳值應該發生在圖中的最低點。

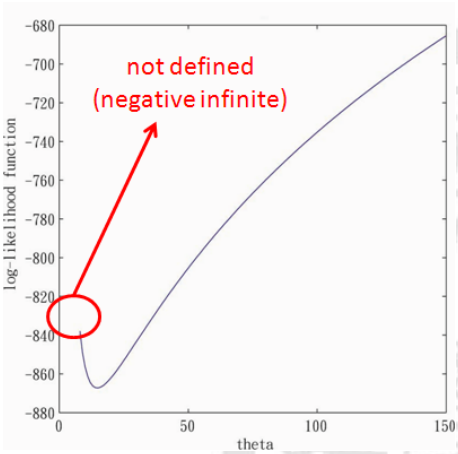
將不同的樣本數與θ的情況下，列出此範例中樣本空間相關函數矩陣的行列式值如表3.2。由表3.2可以觀察到樣本空間相關函數矩陣之行列式值是一個極端數值(Extreme Number)，然而在處理此種極端數值時，往往需要考慮到電腦程式語言所能儲存的資料型態，而本論文所使用的程式語言MATLAB，提供雙精確度(Double)變數資料型態，能儲存的最大值約爲10^{300}左右，一旦數值超出範圍，則程式自動會判斷爲無限大，或是一個沒有意義的數值，因爲這些數值的計算還必須對照樣本點的函數值，所以我們並確定這種情形會在EGO演算法的何時發生，因此此種現象是使用最大似然推估法需要注意的地方。



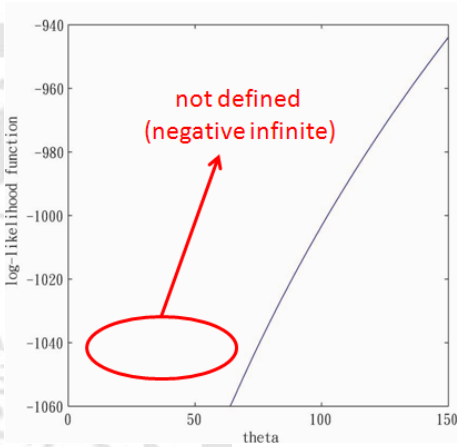
(a) 40個樣本



(b) 60個樣本



(c) 80個樣本



(d) 100個樣本

圖 3.17: 不同樣本數之似然函數與θ對照圖

表 3.2: 樣本空間相關函數矩陣之行列式值

θ	樣本數目				
	20	40	60	80	100
1	7.24×10^{-70}	1.85×10^{-171}	2.07×10^{-282}	0	0
2	4.12×10^{-63}	1.18×10^{-158}	1.41×10^{-263}	0	0
3	5.63×10^{-59}	5.69×10^{-151}	2.30×10^{-252}	0	0
4	6.13×10^{-56}	2.07×10^{-145}	2.69×10^{-244}	0	0
5	1.62×10^{-53}	5.03×10^{-141}	5.77×10^{-238}	0	0

使用最大似然推估法進行Kriging模型擬合時，除了上述之問題外，根據文獻 [35]提到以下幾個問題：

- 多峰值(Multimodality)
- 長山脊(Long Ridges)
- 不良條件(Ill-Condition)

其中第一點多峰值與第二點長山脊，指的是自然對數似然函數(Log-Likelihood Function)其數值本身存在著許多峰值和平緩的平面，依照過去的經驗，此自然對數似然函數通常是一個單調性函數，可參考 [37]，而最早觀察到在高斯歷程中的共變異數，其自然對數似然函數具有這兩個問題的文獻爲 [47]，而文獻中用來說明具有多峰值和長山脊性質如下圖3.18；而第三點不良條件，指的是方程式(3.17)在電腦數值運算上的不穩定現象，可參考 [48]。

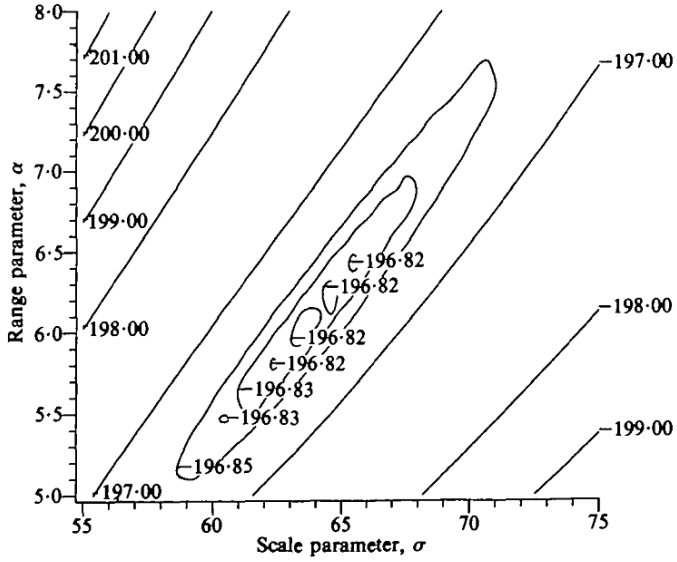


圖 3.18: Warenw et al.說明自然對數似然函數具有多峰值和長山脊圖

目標函數的多峰值與長山脊的性質，將會造成進行最大化時出現演算上的困擾，這使得我們無法使用一些局部最佳化演算法來進行演算，如梯度演算法，而必須利用全域最佳化演算法，如基因演算法、模擬退火演算法，這會造成在擬合時需要花費較多的時間，特別是在使用最大似然嶗龜法且問題具有較多的維度時，擬合的過程會佔去EGO演算法大多數的時間，可參考文獻 [33,49]所提到的現象。

因此演算程式 $superEGO$ 爲了避免耗費過多的時間，選擇在每十次迭代過程中才進行一次的參數擬合；而在文獻 [49]中，歸納出一個演算策略與條件，以避免在每次迭代都進行參數擬合；然而不管是哪一種方法，其目的都是在避免對Kriging模型進行參數擬合的動作，其原因在於對於EGO演算法來說，Kriging模型只是要提供下一個樣本點位置，替演算法決定下一個位置應該往何處搜尋，因此擬合完的Kriging模型是否與實際函數能作到誤差最小反而不是這麼重要，然而一個準確的Kriging模型能夠提供演算法有較快搜索到答案的機會，但沿用著之前所擬合的結果卻可以節省時間，所以事實上這是一種折衷的作法。

本論文程式中參考文獻 [50]的方法，選擇使用美國密西根大學最佳化實驗室所撰寫的DIRECT演算法，演算程式 $UMDIRECT$ 來對Kriging模型進行參數擬合。

在以上考量後，考慮到我們所要處理的問題是最大可靠度問題，因此勢必會對拘束條件的極限狀態(Limite State)作許多取樣，所以本論文程式對於Kriging模型的參數擬合的預設方法爲最小變異圖誤差。

第 四 章　取樣指標 (Infill Sampling Critera)

4.1　常用取樣指標 (Common Infill Sampling Criterion)

取樣指標是EGO演算法中用來決定下一個取樣位置的方法，其決定的方式就是根據某種取樣指標函數進行最佳化，因此依照不同目的所撰寫的取樣指標，會導致演算法取樣至截然不同的位置；在第二章提到，EGO演算法是貝氏分析最佳化演算法的其中一種，而能夠在每次迭代過程中選用不同指標是此種演算法的另一個特色，以下將簡單介紹數個取樣指標。

4.1.1　Kushner

由於所謂的演算法都是要將目標函數最小化或最大化，因此取樣指標很自然的會與目標函數有關，最早使用的取樣指標可參考1964年的文獻 [16]，此文獻中所使用的統計模型是韋納過程，可參考 [51]，如果統計模型是Kriging模型，則改爲方程式(4.1)，我們稱之爲Kushner指標。

$$
\begin{aligned}
ISC_{\text{Kushner}}(\mathbf{x}) &= P_{\text{r}}(\hat{f}(\mathbf{x}) < f_{min} - \epsilon) \\
&= \Phi\left(\frac{(f_{min} - \epsilon) - \hat{f}(\mathbf{x})}{\hat{\sigma}_f(\mathbf{x})}\right)
\end{aligned}
\tag{4.1}
$$

上式中，$\hat{f}$爲Kriging模型根據取樣所建立的目標函數預測值平均值，而$\hat{\sigma}_f$是對應的標準差，f_{min}是現有的取樣點中函數值最小的量值，ϵ是一個使用者自訂的參數，根據其數學式來說明Kushner指標的意義，即爲Kriging模型的預測值比目前最好的答案減去ϵ還要小的機率爲多少。而Kushner指標藉由調整ϵ來決定此次迭代趨向於尋找全域(Global)或是局部(Local)的取樣，例如$0.01 \times f_{min}$或$0.001 \times f_{min}$，兩者的不同可以參考圖4.1，圖4.2。

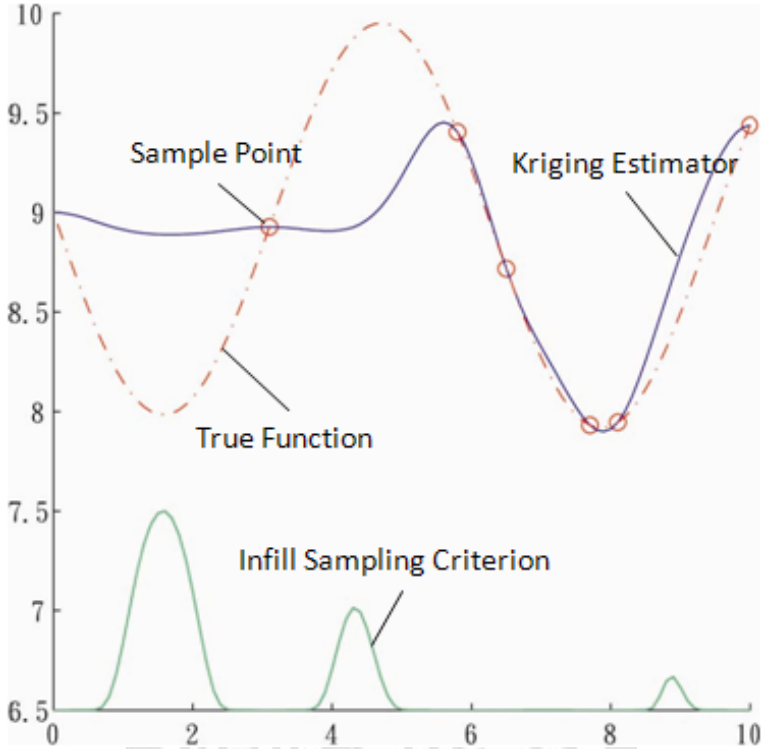


圖 4.1: $\epsilon = 0.1 \times f_{min}$之Kushner指標圖

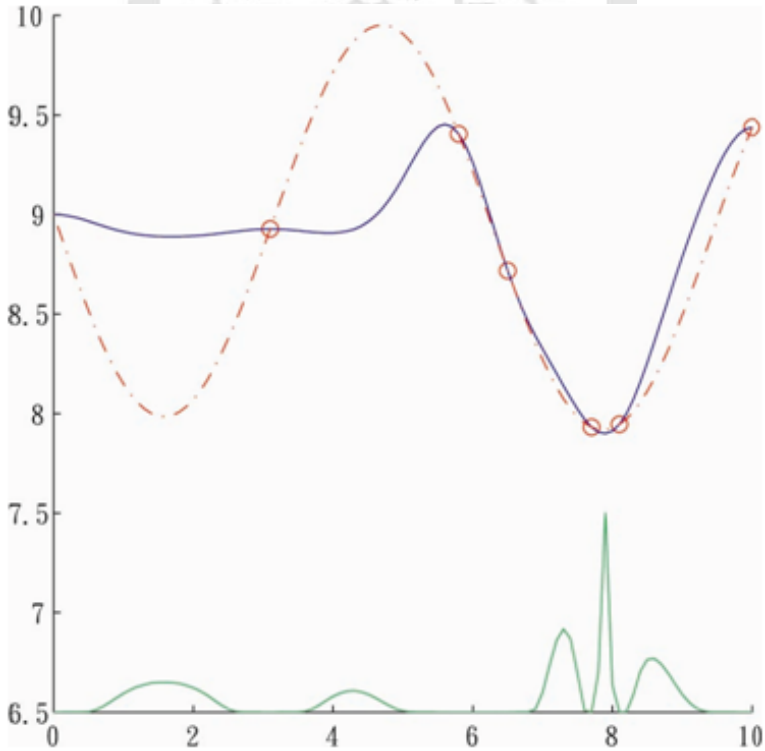


圖 4.2: $\epsilon = 0.01 \times f_{min}$之Kushner指標圖

其中上方虛線爲原函數，o爲樣本點，上方實線爲Kriging模型的預測平均值，底部的實線爲Kushner指標正規化後的數值。演算法下一個取樣的位置，即爲圖4.1中底部實線最高的位置，由這兩張圖我們可以看出，當ϵ越大時，演算法傾向搜尋全域的答案，當ϵ越小，則演算法傾向搜尋局部的答案。

4.1.2 Expected Improvement

由於在Kushner指標中，如何選擇一個適當的ϵ值是一件困難的事，因此才有後來各種取樣指標的發展。如同第二章提到，EGO演算法所使用的取樣指標爲Expected Improvement(EI)，姑且將其翻譯爲改善期望值指標，其公式推導可參考 [34]，首先我們先定義何謂改善，如方程式(4.2)。

$$I = \max\left\{0, f_{min} - \hat{f}(\mathbf{x})\right\} \tag{4.2}$$

依照其數學式，改善的量值定義爲，現有最小的取樣點量值減去Kriging模型的預測平均值與零之間的大值，因此除非Kriging模型的預測平均值比現有最小的取樣點量值還小，否則就是零改善，然而我們並不是只用此方程式來當作取樣指標，而是取改善量值的期望值，如下方程式(4.3)，我們稱此式爲改善期望值(Expected Improvement)。

$$ISC_{\mathrm{EI}}(\mathbf{x}) = E(I) = \begin{cases} (f_{min} - \hat{f}(\mathbf{x}))\Phi(k) + \hat{\sigma}_f(\mathbf{x})\phi(k), & \text{if } \hat{\sigma}_f > 0 \\ 0, & \text{if } \hat{\sigma}_f = 0 \end{cases} \tag{4.3}$$

$$k = \frac{f_{min} - \hat{f}(\mathbf{x})}{\sigma_{\hat{f}}(\mathbf{x})}$$

將方程式(4.2)改寫成方程式(4.4)如下，其中g爲一個非負整數，我們稱此式爲改善量值的一般式。

$$I^g = \max\left\{0, (f_{min} - \hat{f}(\mathbf{x}))^g\right\} \tag{4.4}$$

對應此一般式的期望值如下方程式(4.5)，我們稱此式爲改善期望值一般式(Generalized Expected Improvement)。

$$
\begin{aligned}
ISC_{\mathrm{GEI}} = E(I^g) = \hat{\sigma}_f^g \sum_{i=0}^{g} (-1)^i \left(\frac{g!}{i!(g-i)!} \right) k^{g-i} T_i \qquad (4.5)\\
T_0 = \Phi(k), \quad T_1 = -\phi(k)\\
T_i = -\phi(k) k^{i-1} + (i-1) T_{i-2}\\
k = \frac{f_{min} - \hat{f}}{\hat{\sigma}_f}
\end{aligned}
$$

改善期望值一般式藉由改變g的量值，來決定演算法搜索的取向，當$g = 1$時，則改善期望值一般式將退化成方程式(4.3)，關於g值對改善期望值指標的影響可參考下圖4.3、圖4.4。

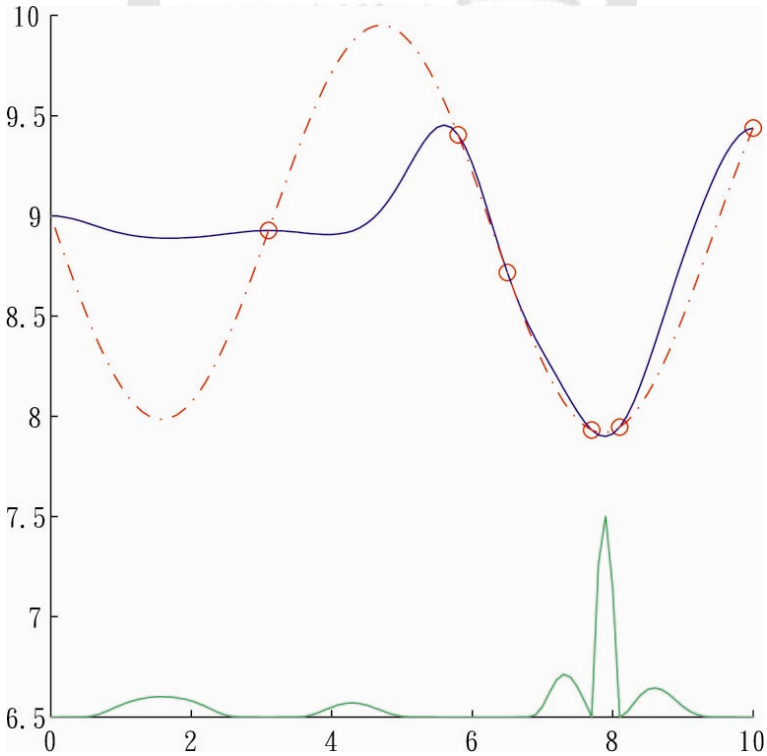


圖 4.3: $g = 1$之改善期望值指標圖

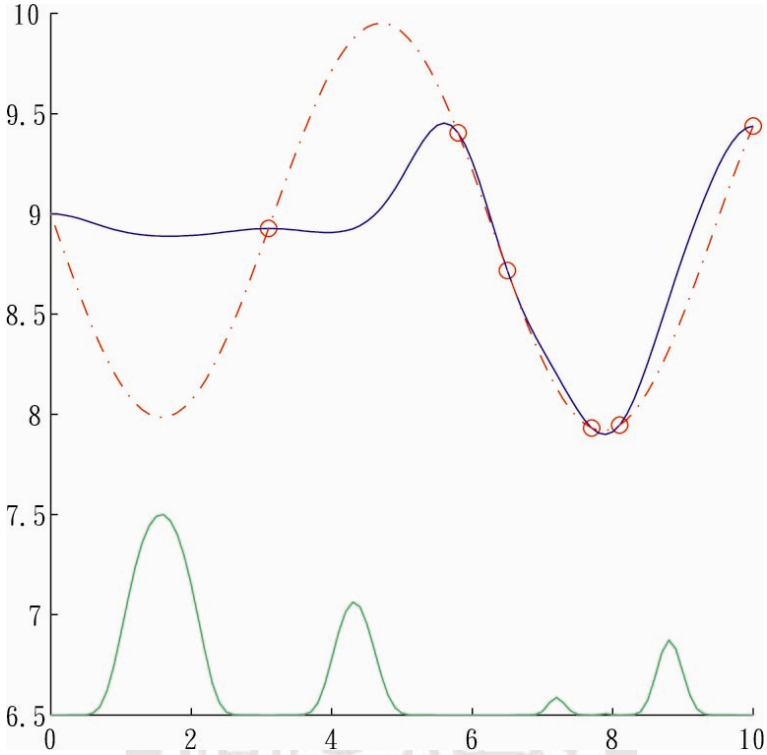


圖 4.4: $g = 5$之改善期望值指標圖

由這兩張圖我們可以看出，當g值越大時，演算法傾向搜尋全域的答案，當g越小，則演算法傾向搜尋局部的答案。而在文獻 [44]中，將此指標與模擬回火演算法的概念類比，制定在每個不同迴圈中使用不同的g值，並稱之爲Cooling指標。

4.1.3 Kriging Estimator

針對搜尋目標函數的指標還有很多，然而最簡單的指標形式其實是Kriging模型的預測平均值，也就是$\hat{f}$。

由之前的Kushner指標與改善期望值指標都可以發現，其本身都是具有許多峰值的函數，並擁有數值爲零的一個很大的平面，在本章一開始提到，EGO演算法中用來決定下一個取樣位置的方法是根據某種取樣指標函數進行最佳化，然而上述這些函數的特性，都有可能會增加最佳化演算法在演算時的困難，而只使用Kriging模型的預測平均值作爲取樣指標的一個好處是，此指標在設計空間中的形狀取決於問題本身，因此在大部分的情況下會具有比較好的演算效果。

另一個使用Kriging模型的預測平均值作爲取樣指標的好處是，此指標只使用一個

量值$\hat{f}$，而此量值與取樣值f的性質重疊，兩者代表著同一個東西，只是一個是眞實試驗所得到的數據，另一個是Kriging模型得到的預測值，此特性所具有的好處是，這使得EGO演算法不再侷限於解決黑盒子函數，而是一個能與一般數學問題結合的演算法，這也是演算程式$superEGO$的將原本的EGO演算法推廣到解決各種不同問題的方法，詳細可參考 [52]。

本論文程式中，用來針對目標函數得到更好的取樣方法，以此取樣指標爲預設值。

4.1.4 Maximum Varianve

以上介紹的取樣指標，都是針對目標函數所製作，然而另一種簡單的指標是Kriging模型的預測值變異數，也就是$\hat{\sigma}$，其修正過程如下圖4.5。

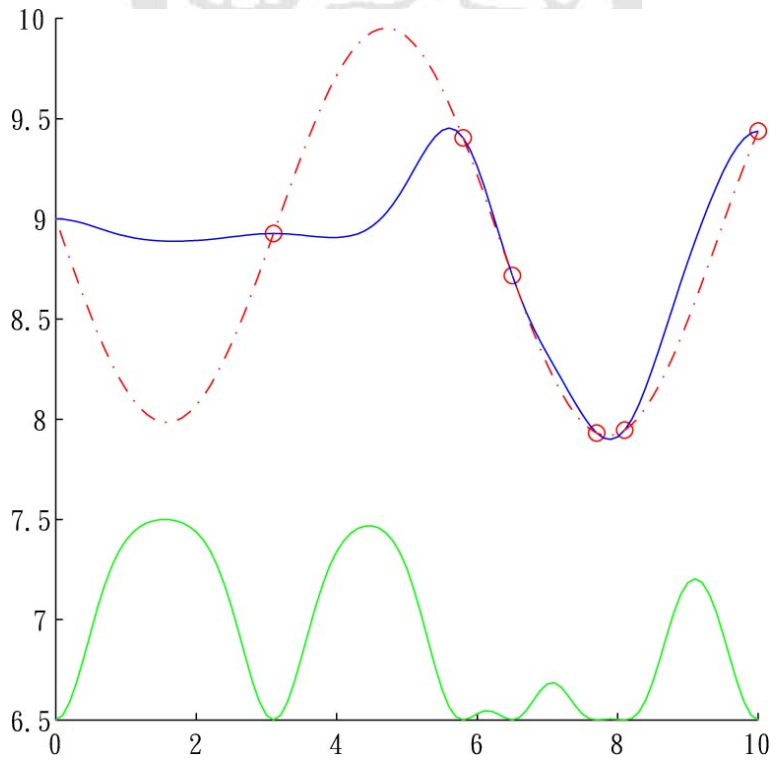
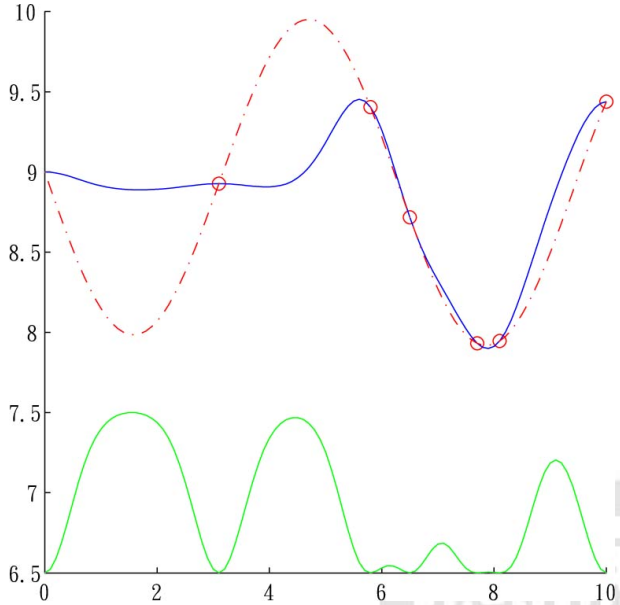
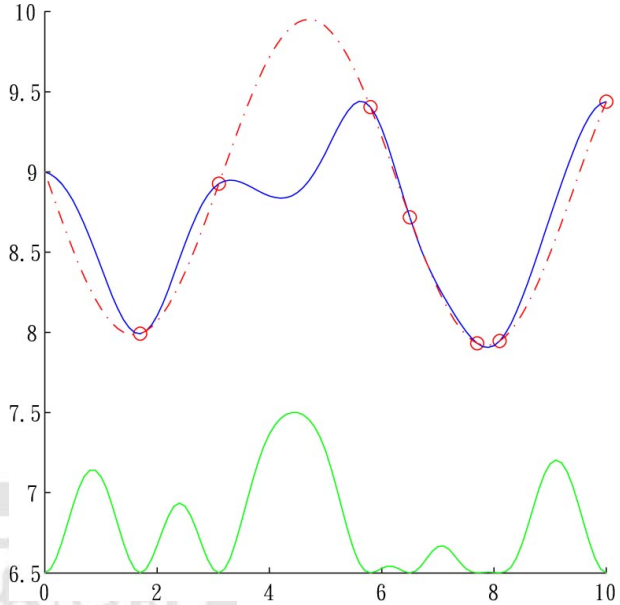


圖 4.5: Maximum Variance指標圖

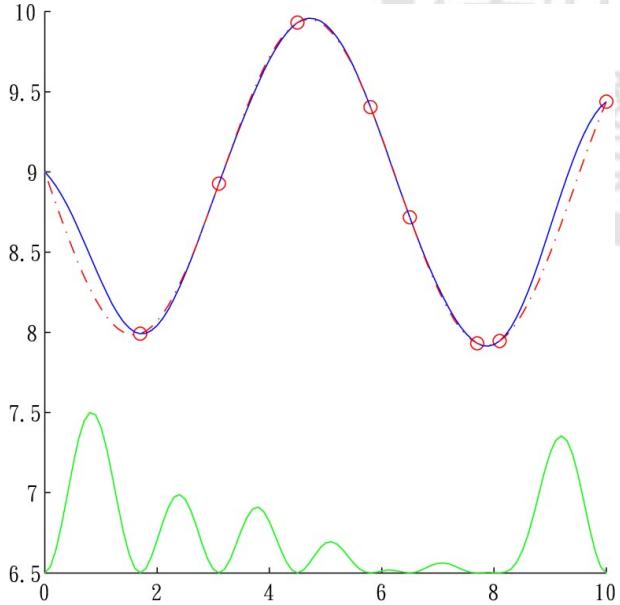
我們稱此指標爲最大變異數指標(Maximum Varianve)，此指標是最常用來修正Kriging模型的指標，如圖4.6。



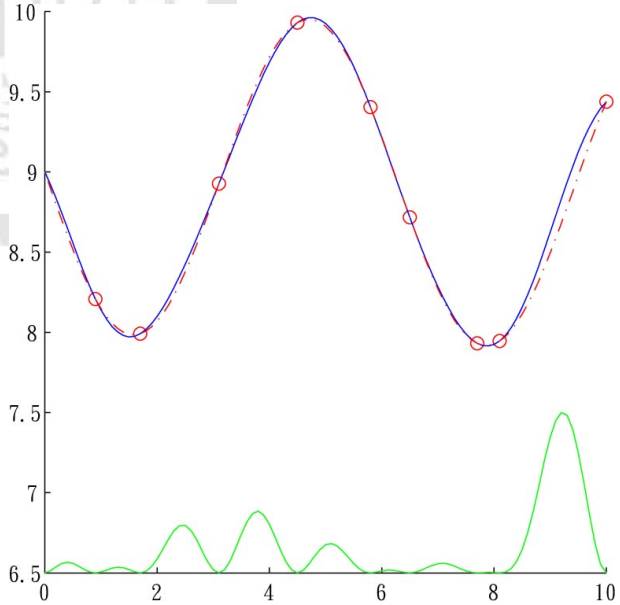
(a) 7個取樣



(b) 8個取樣



(c) 9個取樣



(d) 10個取樣

圖 4.6: 利用最大變異數指標修正函數過程圖

4.1.5 Probability of Feasibility

能夠依照不同目的而撰寫取樣指標，是EGO演算法具有彈性的地方，以下將介紹兩種較爲特殊的取樣指標，而這兩種指標所針對的對象爲拘束條件。

最佳化演算法往往必須考慮拘束條件，一個滿足拘束條件的答案我們稱之爲可行解，可行解所構成的變數空間我們稱之爲可行解空間，一個不滿足拘束條件的答案，無論其目標函數值的表現多好，其本身仍舊不是解，然而對於實際的問題來說，黑盒子函數本身並不一定來自於目標函數，也很有可能是來自於拘束條件，也就是說拘束條件量值的取的，也有可能是必須通過黑盒子函數的計算才能取得，因此如何有效率的解決拘束條件是黑盒子函數也是EGO演算法的一個重要課題，需要探討的情況分爲兩種，分別是：

- 目前取樣尚無可行解。
- 如何針對非可行解空間搜索可行解。

針對第一個問題，由於實際的取樣才能作爲演算結果的答案，所以目前所有的取樣尚無可行解時，等同於目前所有取樣中沒有一個可以作爲尚可接受的答案，因此取樣指標的主要目標是要試圖找到一個有可行解的位置，參考文獻 [18]，我們稱此指標爲可行解機率(Probability of Feasibility)，其數學式如下方程式(4.6)，當有多個拘束條件時，可改爲方程式(4.7)。

$$ISC_{\text{Feasibility}}(\mathbf{x}) = \Phi\left(\frac{0-\hat{g}(\mathbf{x})}{\hat{\sigma}_g(\mathbf{x})}\right) \tag{4.6}$$

$$= \prod_{i=1}^{k} \Phi\left(\frac{0-\hat{g}_i(\mathbf{x})}{\hat{\sigma}_{g_i}(\mathbf{x})}\right) \tag{4.7}$$

以Branin範例 [21]說明：

$$g(\mathbf{x}) = \left(x_2 - \frac{5.1}{4\pi^2}x_1^2 + \frac{5}{\pi}x_1 - 6\right)^2 + 10\left(1 - \frac{1}{8\pi}\right)\cos x_1 + 7 \leq 0$$

其$g = 0$與取樣點位置如圖4.7，其中黑色實線爲$g = 0$位置，實線內的範圍爲可行解空間，紅色o爲樣本位置，可以注意到目前取樣中並未有可行解，將其可行解機率指標繪製，如圖4.8。

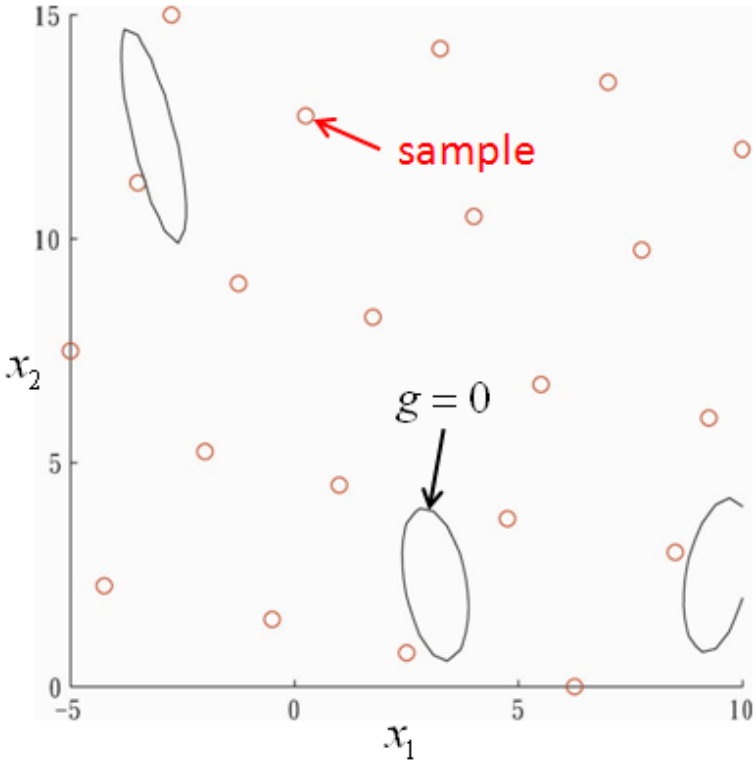


圖 4.7: 取樣與$g = 0$位置圖

由圖上觀察可知，此取樣指標的函數行爲與可行解空間的範圍大致上符合，因此在目前沒有取樣是可行解的情況下，我們可以利用此取樣指標來找到可行解的位置。然而依照此指標選取樣本後發現，可行解機率指標在目前取樣已存在可行解的情況下，會開始在已知的可行解區域內進行取樣，如圖4.9，與圖4.7比較，圖4.9在左上方可行解區內多了三個樣本，圖形以x作爲標示，但是其對應的可行解機率指標如圖4.10，卻在已知的可行解區域有最高值。

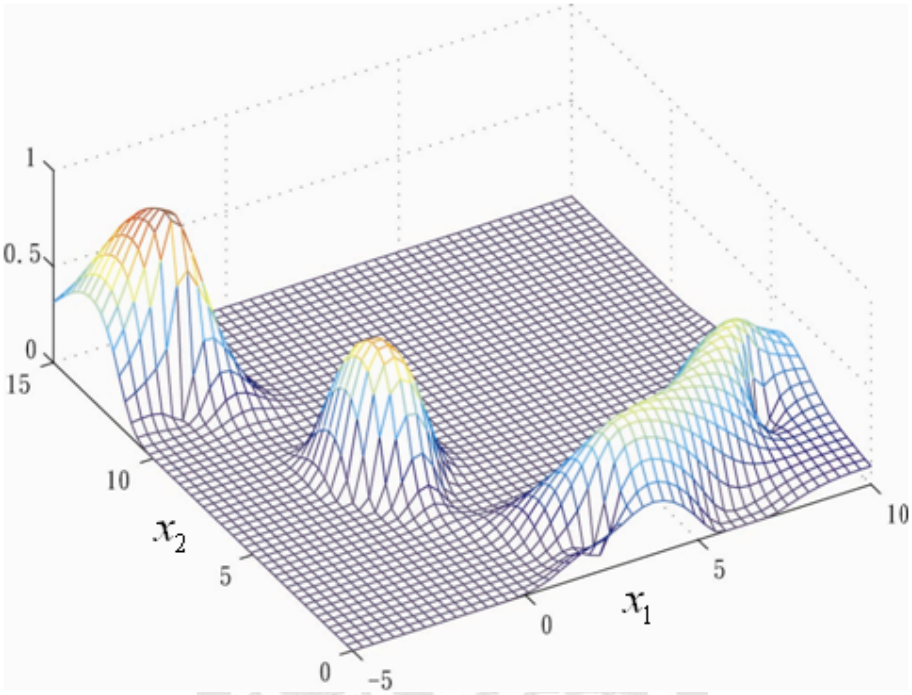


圖 4.8: 可行解機率指標圖

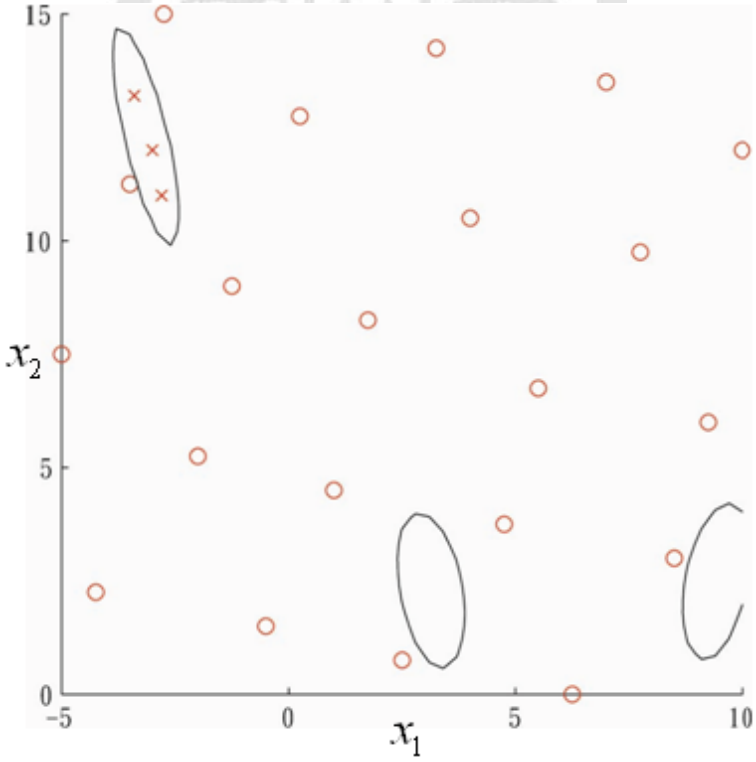


圖 4.9: 取樣與$g = 0$位置以及新增取樣位置圖

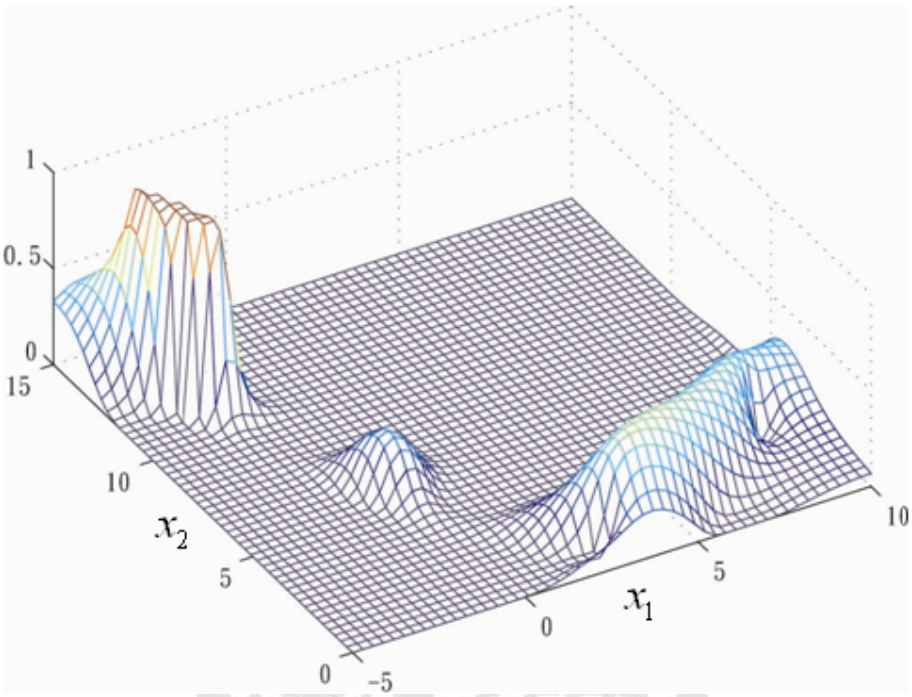


圖 4.10: 新增取樣後之可行解機率指標圖

此現象即爲尚未探討的第二個現象，如何針對非可行解空間搜索可行解，在下一個指標說明。

4.1.6 Tunneling

同樣參考文獻 [18]，針對非可行解空間搜索可行解，我們稱之爲隧道指標(Tunneling)，其數學式如下方程式4.8。

$$ISC_{\text{Tunneling}}(\mathbf{x}) = \Phi\left(\frac{0-\hat{g}(\mathbf{x})}{\hat{\sigma}_g(\mathbf{x})}\right)\cdot\left[\min\left(\frac{\|\mathbf{x}_{\text{f}}-\mathbf{x}\|}{\|\mathbf{x}_{\text{UB}}-\mathbf{x}_{\text{LB}}\|}\right)\right] \tag{4.8}$$

上式中，前面的部份與可行解機率指標相同，不一樣的地方在於後面多乘了一個與樣本之間距離有關之函數，其中$\mathbf{x}_{\text{f}}$代表可行解樣本的位置，$\mathbf{x}_{\text{UB}}$與$\mathbf{x}_{\text{LB}}$分別代表著變數空間的上限與下限，由於變數空間的上限與下限之間的距離是固定的，所以在此取樣指標中，等於是可行解機率指標多乘上取樣與任何一個可行解之間的距離中的最小值，這會造成此取樣指標會傾向於取樣在遠離現有已知可行解樣本，延續上一個範例如圖4.9，將此時的隧道指標繪製，如圖4.11。

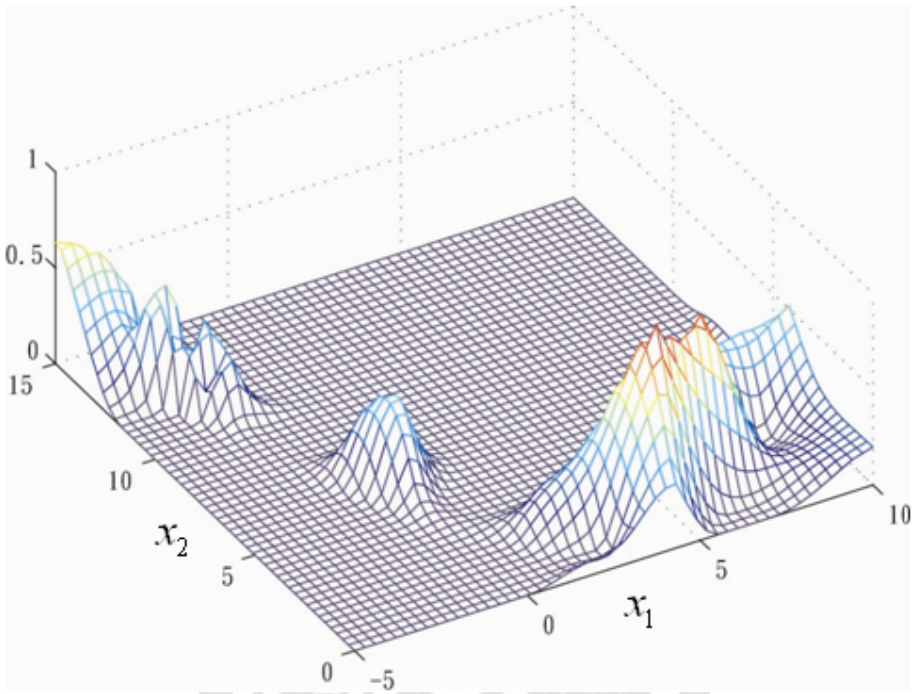


圖 4.11: 新增取樣後之隧道指標圖

由圖中可以觀察到，此時的取樣指標已不再局限於已知的可行解空間，也正因爲此指標有如在各個分離的可行解空間中挖取隧道一般，所以才稱之爲隧道指標，當有多個拘束條件時，隧道指標可改爲方程式(4.9)。

$$ISC_{\text{Tunneling}}(\mathbf{x}) = \left[\prod_{i=1}^{k} \Phi\left(\frac{0-\hat{g}_i(\mathbf{x})}{\hat{\sigma}_{g_i}(\mathbf{x})}\right)\right] \cdot \left[\min\left(\frac{\|\mathbf{x}_{\text{f}}-\mathbf{x}\|}{\|\mathbf{x}_{\text{UB}}-\mathbf{x}_{\text{LB}}\|}\right)\right] \tag{4.9}$$

至此可以得到一個結論，就是當目前取樣沒有任何可行解時，我們必須利用可行解機率指標來尋找可行解空間，一旦有取樣是可行解時，我們必須利用隧道指標來進行對非可行解空間中可行解空間的探索。

4.2 取樣指標分類 (Classification of ISC)

本章只介紹了一些常用的取樣指標來說明取樣指標的概念，然而文獻中還有其他取樣指標可以選擇，例如Lower Confidence Bounding Function [53]、Locating the Threshold Extreme、Locating the Regional Extreme、Minimizing Surprises [54]、Interpolating Variance [55]、Clustered Multiple Generalized Expected Improvement [56]...等。

取樣指標在制定上並不困難，端看使用者希望此指標要達成什麼目的，因此在文獻中許多指標都有功能性上的重複，主要將其分類爲三種：針對目標函數最佳值搜索、減少Kriging模型誤差、具有特殊目的，如下表4.1、表4.2、表4.3。

表 4.1: 最佳值搜索類型取樣指標表

指標名稱	作者	年份
Kriging Estimator	G. Mathron	1963
Kushner	H. J. Kushner	1964
Lower Bounding Function	D. D. Cox, S. John	1992
Locating the Threshold Extreme	A. G. Watson, R. J. Barnes	1995
Locating the Regional Extreme	A. G. Watson, R. J. Barnes	1995
Expected Improvement	M. Schonlau, W. J. Welch, D. R. Jones	1997
Generalized Expected Improvement	M. Schonlau, W. J. Welch, D. R. Jones	1997
Cooling	M.J. Sasena	2002
CMG Expected Improvement	W. Ponweiser, T. Wagner, M. Vincze	2008

表 4.2: 減少誤差類型取樣指標表

指標名稱	作者	年份
Maximum Variance	G. Mathron	1963
Minimizing Surprises	A. G. Watson, R. J. Barnes	1995
Interpolating Variance	J. K. Yamamoto	2000

表 4.3: 特殊目的類型取樣指標表

指標名稱	作者	年份
Probability of Feasibility	M. J. Sasena	2002
Tunneling	M. J. Sasena	2002
Reduce Constraint Slackness	M. J. Sasena	2002

上表列出之取樣指標可能並不完善，而在此我們並不比較各種同類型的取樣指標的優劣，我們只將各種取樣指標作初步的分類，並附上作者與年份，提供本論文程式使用者在使用上的一個參考依據。

4.3 聯合極限狀態取樣指標 (Sampling on Joint Limit State)

極限狀態(Limit State)指的是拘束條件$g = 0$的位置，由於我們要處理的問題是針對黑盒子函數拘束條件來計算可靠度問題，然而要針對黑盒子函數來計算可靠度，其本身就是一件困難的事，最大的困難在於我們無法透過多次計算函數值來取得眞實拘束條件的資訊。

在第二章可靠度分析法的內容提到，無論是哪一種方法得到可靠度都需要對函數進行額外的計算，例如：蒙地卡羅法需要取十到百萬個函數值才能估計可靠度，一階可靠度方法需要計算函數的梯度或是經過最佳化的迭代取得最有可能破壞點(MPP)才能估計可靠度，如果計算一次黑盒子函數需要一個小時，那麼可靠度最佳化設計就將是一件以曠日費時的事情，所以這些方法都無法直接使用在黑盒子函數問題。

因此過去解決此問題的作法都是先對黑盒子函數取樣，然後再利用這些樣本來製作反應曲面(Response Surface、Surrogate、Metamodel)以取代此黑盒子函數，之後根據此反應曲面進行可靠度最佳化設計，可參考文獻 [30,57]。然而針對此種方法我們不禁要提出疑問，就是如何確定反應曲面與黑盒子函數之間的誤差可以忽略？即使反應曲面與黑盒子函數之間的誤差可以忽略，又如何可以確定此誤差不會影響可靠度的計算？

將Camelback函數 [2]平移一個常數項，使其可行解空間相連：

$$g(\mathbf{x}) = -0.1 + \left(4 - 2.1x_1^2 + \frac{x_1^4}{3}\right)x_1^2 + x_1x_2 + \left(-4 + 4x_2^2\right)x_2^2 \leq 0$$

此範例與方程式(3.28)相差一個常數項，使得極限狀態的位置稍有不一樣，可參考下圖4.12。

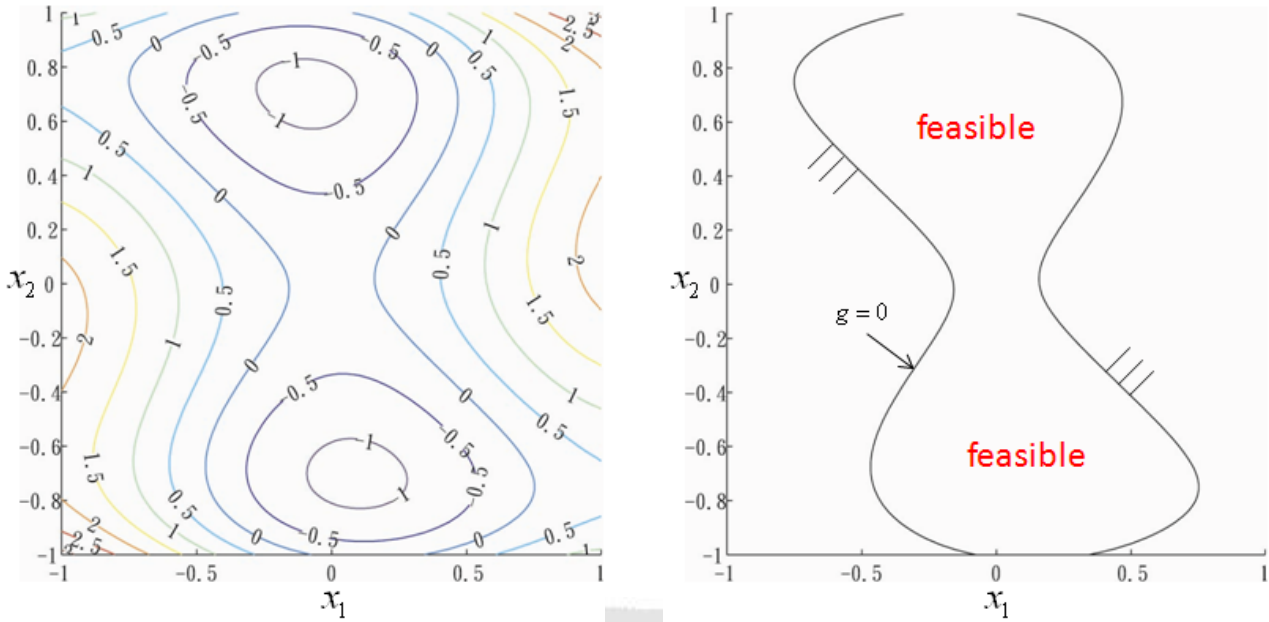


圖 4.12: Camelback函數與其可行解區域位置圖

上圖4.12左側是此函數的等高線圖，右側圖中的黑線是此函數的極限狀態，根據此函數分別進行兩次不同樣本內容的取樣，將取樣後所重建的Kriging模型以及其對應的極限狀態繪製成圖，如圖4.13、圖4.14，兩次取樣的樣本數一樣多，但是分別建立的Kriging模型之等高線圖與極限狀態圖皆不一樣，而兩次取樣的主要差異是，第二次取樣中有較多的樣本取在接近極限狀態的位置。

RMS(Root Mean Square Error)是一種拿來判斷反應曲面與實際函數誤差的指標，如下列方程式(4.10)，其中N是計算此指標所取的樣本。

$$RMS = \frac{1}{N}\sqrt{\sum_{i=1}^{N} \hat{g}(\mathbf{x}_i) - g(\mathbf{x}_i)} \tag{4.10}$$

我們對此範例兩次取樣的結果，分別取900個網格化的位置，也就是$N = 900$，來計算這兩個Kriging模型的RMS，其結果如表4.4。

表 4.4: 兩次取樣RMS表

	第一次取樣	第二次取樣
RMS	0.0659	0.0663

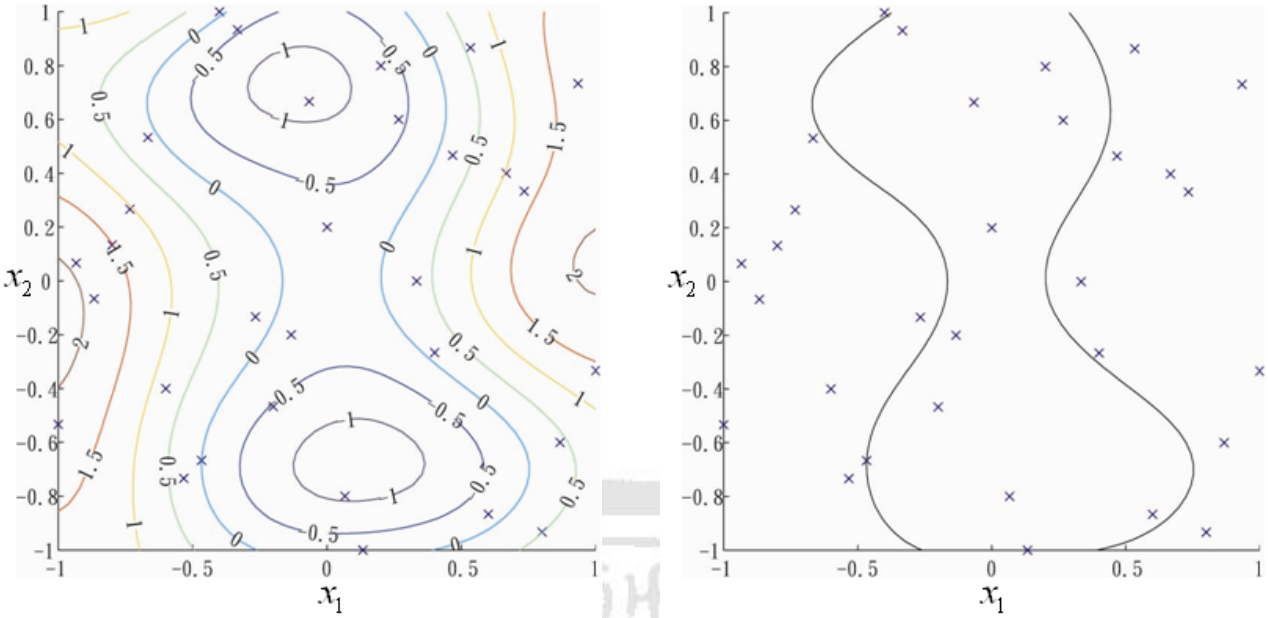


圖 4.13: 第一次取樣結果之Kriging模型與其可行解區域位置圖

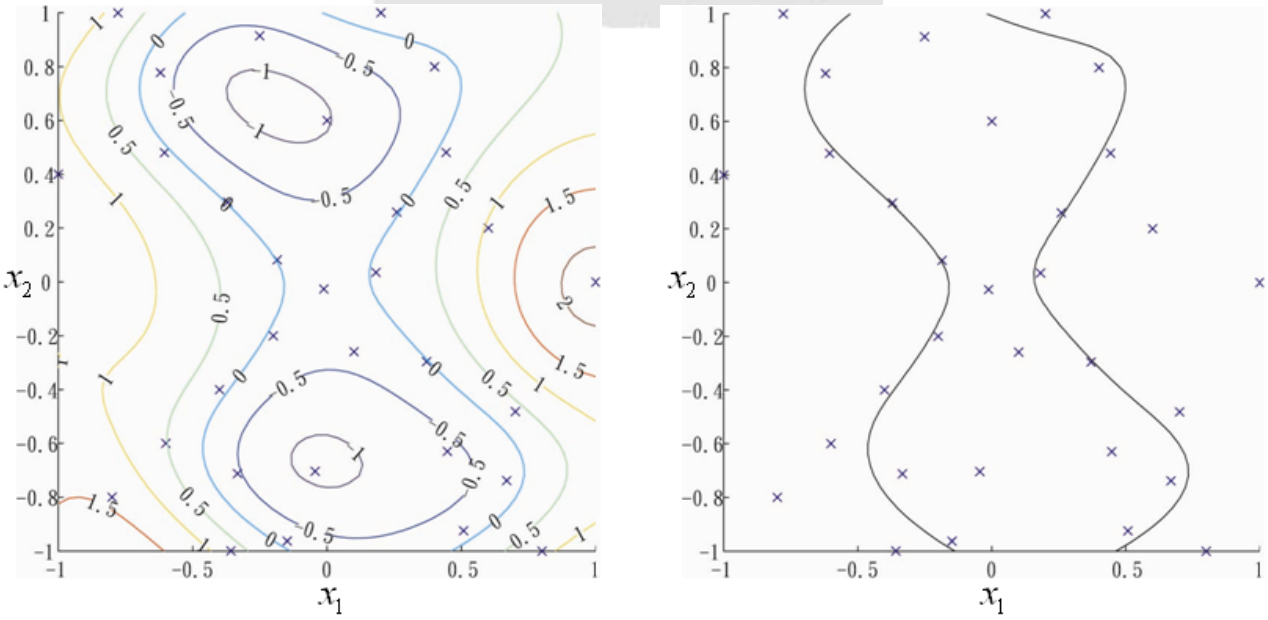


圖 4.14: 第二次取樣結果之Kriging模型與其可行解區域位置圖

由RMS的結果可以知道，第一次取樣的Kriging模型具有比較好的整體預測效果，我們將其原因歸於第一次取樣具有較好的樣本分散程度。

現在考慮在設計點$(0,0)$位置有一個高斯分佈隨機變數$\mathbf{X}$，如方程式(4.11)。

$$\begin{aligned} \mathbf{X} &= (X_1, X_2) \\ X_1 &\sim N(0, 0.5) \\ X_2 &\sim N(0, 0.3) \end{aligned} \tag{4.11}$$

我們使用蒙地卡羅法在$(0,0)$取樣十萬點，分別對原函數、第一次取樣Kriging模型、第二次取樣Kriging模型，進行可靠度的計算，計算的結果如下表4.5，由可靠度測試結果可以得知，第二次取樣的Kriging模型在計算可靠度上能有比較準確的結果。

表 4.5: 蒙地卡羅法測試結果表

	原函數	第一次取樣	第二次取樣
可靠度	42.95%	43.80%	43.01%

至此，我們可以得到一個結論，在針對一個黑盒子函數計算可靠度時，最重要的一件事不是反應曲面與原函數的整體誤差爲多少，而是所製作的反應曲面，其曲面上的極限狀態是否能與原函數的極限狀態相符合，只要此判斷可行解與非可行解的界線能夠清楚，曲面上其他位置的數值爲多少就不是這麼重要了。

因此我們需要一個能夠試圖在極限狀態上取樣的指標，然而在過去的文獻中並沒有符合我們需求的指標，在參考隧道指標的概念後，我們將此指標的數學式定義如下列方程式(4.12)，我們稱之爲極限狀態取樣指標(Sampling on Limit State Criterion)。

$$ISC_{\mathrm{SoLS}}(\mathbf{x}) = \delta(\mathbf{x}) \cdot \left[\min\left(\frac{\|\mathbf{x}_{\mathrm{s}} - \mathbf{x}\|}{\|\mathbf{x}_{\mathrm{UB}} - \mathbf{x}_{\mathrm{LB}}\|}\right)\right] \tag{4.12}$$

$$\delta(\mathbf{x}) = \begin{cases} 1, & \text{if } \|\hat{g}_i(\mathbf{x})\| \leq \Delta \\ 0, & \text{else} \end{cases}$$

此式中的$\mathbf{x}_s$指的是所有樣本點，Δ爲使用者自定義數；在此指標中提供數值大小的部份爲一個距離的函數，$\delta(\mathbf{x})$的部份只提供一種過濾(Filter)的效果，當Δ數值小的時候，指標結果會靠近Kriging模型上的極限狀態，而且會盡量遠離其他樣本，然而會使指標函數產生峰值的情形，造成最佳化的演算法演算上的困難。

另一種取樣指標是在考慮多個拘束條件下，爲了避免取樣在多餘的極限狀態時所定義的指標，我們稱之爲聯合極限狀態取樣指標(Sampling on Joint Limit State Criterion)，如方程式(4.13)，其中n代表問題的拘束條件總數。

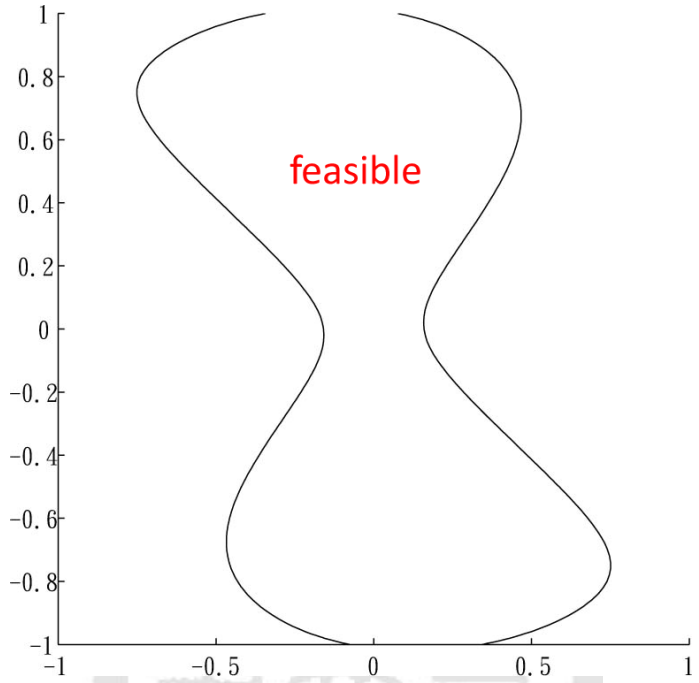
$$
\begin{aligned}
ISC_{\mathrm{SoJLS}}(\mathbf{x}) &= \delta(\mathbf{x}) \cdot \left[\min\left(\frac{\|\mathbf{x}_{\mathrm{s}} - \mathbf{x}\|}{\|\mathbf{x}_{\mathrm{UB}} - \mathbf{x}_{\mathrm{LB}}\|}\right)\right] \\
\delta(\mathbf{x}) &= \begin{cases} 1, & \text{if } \left(\sum_{i=1}^{n} m_i\right) > 0 \\ 0, & \text{else} \end{cases} \\
&\text{for } i = 1...n, \\
m_i(\mathbf{x}) &= \begin{cases} 1, & \text{if } \|\hat{g}_i(\mathbf{x})\| \le \Delta \text{ and } \hat{g}_j(\mathbf{x}) \le \Delta, \text{for } j = 1...n,\ j \ne i \\ 0, & \text{else} \end{cases}
\end{aligned}
\tag{4.13}
$$

以下利用Camelback範例與Baseball Function範例，來觀察本節兩種指標的行爲，其中Baseball Function範例是三個拘束條件凹形可行解空間的問題，如下方程式(4.14)。

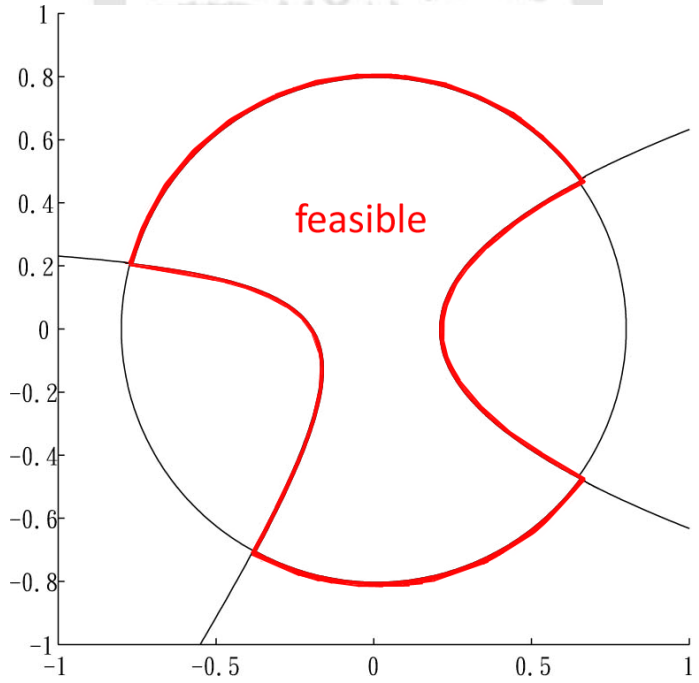
Baseball Function範例：

$$
\begin{aligned}
g_1 &= x_1^2 + x_2^2 - 0.8^2 \le 0 \\
g_2 &= -x_1 + 2x_2^2) + 3x_1x_2 - 0.2 \le 0 \\
g_3 &= x_1 - 2x_2^2 - 0.2 \le 0
\end{aligned}
\tag{4.14}
$$

將兩個範例的極限狀態位置重新繪製，如圖4.15，其中在Baseball Function範例圖中的粗實線爲聯合極限狀態。對此兩個問題，分別使用聯合極限狀態取樣指標來修正問題的極限狀態，其過程如圖4.16、圖4.17。由圖形的演變結果可以看到，我們藉由不斷地在問題的極限狀態上進行取樣，以修正Kriging模型與實際函數的誤差。

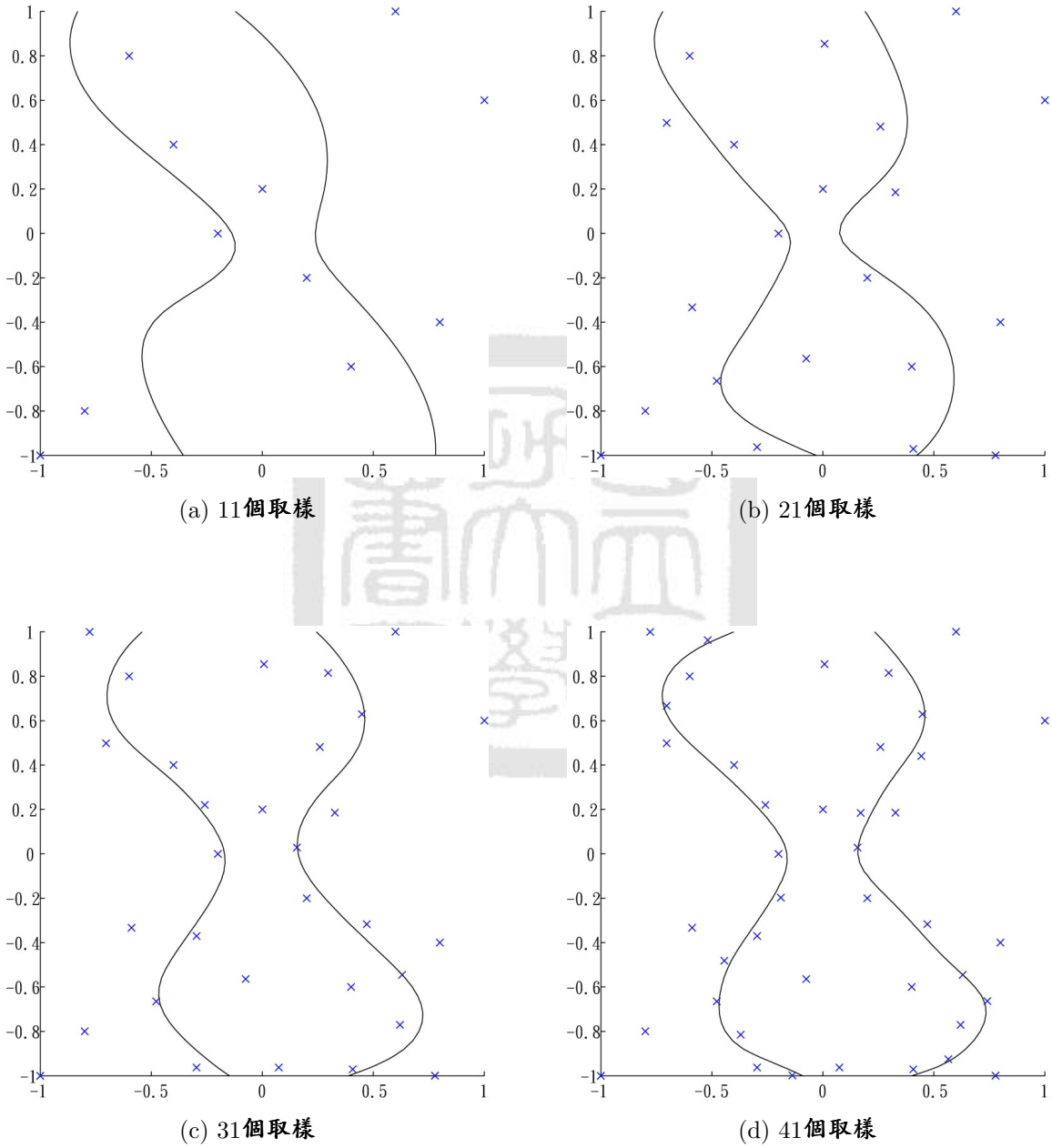


(a) Camelback範例



(b) Baseball Function範例

圖 4.15: 範例極限狀態圖



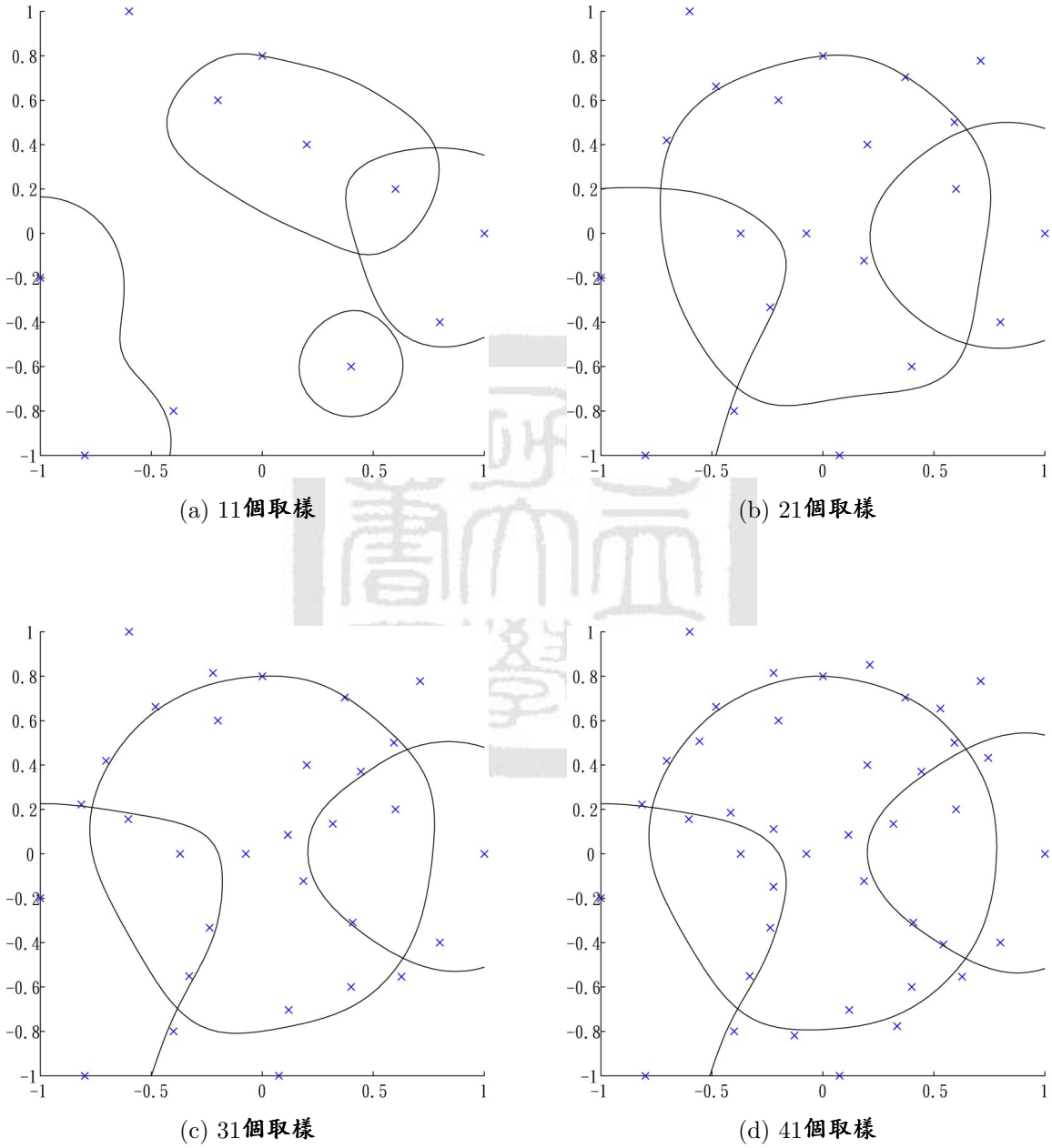
(a) 11個取樣

(b) 21個取樣

(c) 31個取樣

(d) 41個取樣

圖 4.16: 極限狀態取樣指標修正Camelback過程圖



(a) 11個取樣

(b) 21個取樣

(c) 31個取樣

(d) 41個取樣

圖 4.17: 聯合極限狀態取樣指標修正Baseball Function過程圖

第 五 章　演算策略 (Algorithm Strategy)

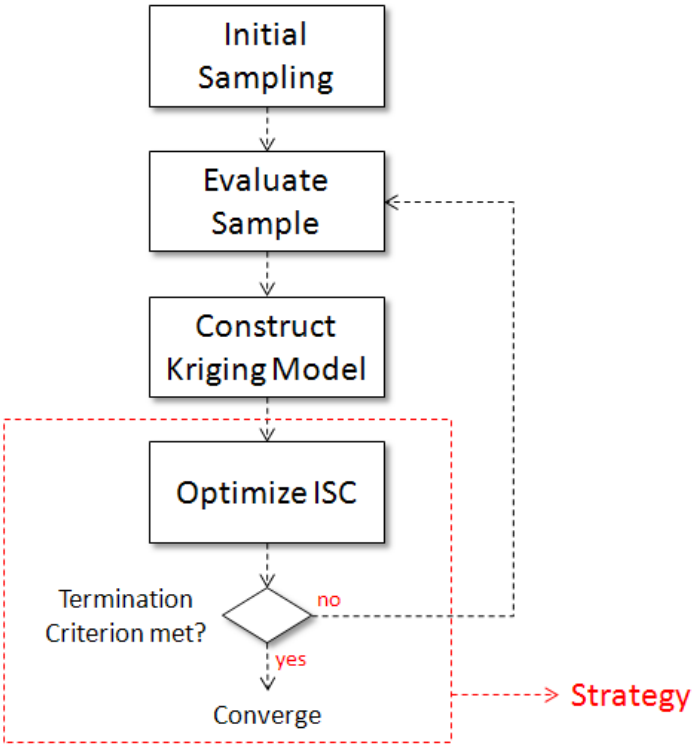


圖 5.1: EGO演算流程圖

如圖5.1，此演算流程與第二章的文獻 [18]中之流程圖2.3沒有太大的差異，這是因爲EGO演算法本身的架構即是如此，然而能夠使EGO演算法對應各種不同問題的關鍵，在於演算策略的制定，如圖5.1中虛線所圈選的部份，其中以如何處理取樣指標最爲關鍵，以下將對文獻中已存在的最佳化演算策略，以及本論文探討的可靠度問題演算策略加以說明。

5.1　最佳化 (Optimization)

此處的最佳化指的是定量變數的最佳化設計，貝氏分析演算法最初發展是針對只有變數上下限的無拘束條件最佳化問題所設計，而在最早的EGO演算法裡使用改善期望值指標(EI)，只需要在每次迭代中選擇不同的改善期望值指標階數，即可達到全域與局部搜

索的效果，請參考第四章，因此對於一個無拘束條件的最佳化問題時，其演算策略如下圖5.2。

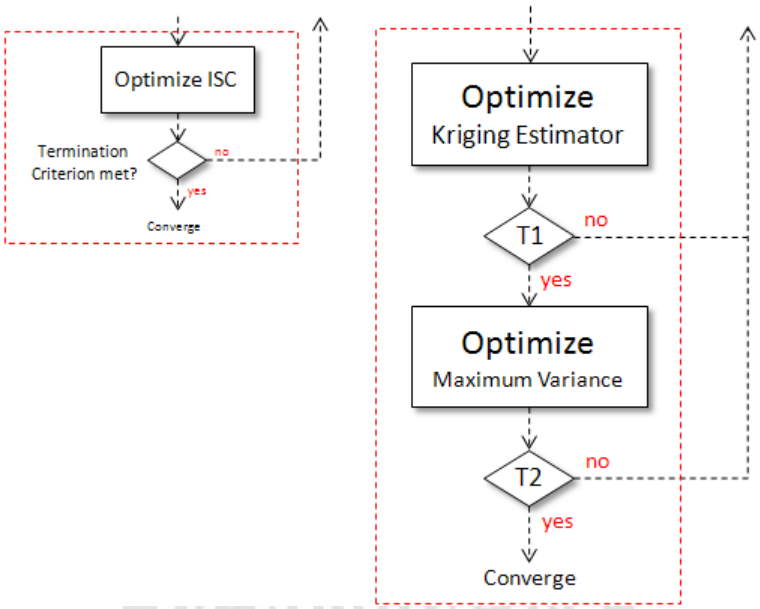


圖 5.2: 無拘束條件演算策略圖

圖5.2中關於T1與T2的部份，代表的是演算流程中的終止條件(Termination)，其意思是指判斷目前情況是否滿足某些條件，來決定演算流程的方向，將上圖中的終止條件的內容列出，如下表5.1。

表 5.1: 無拘束條件演算策略之終止條件

編號	條件內容
1	改善目標函數指標與其他樣本過近
2	最大變異數指標量值小於使用者自訂數

此部份流程圖爲圖5.1中演算策略的部份，其主要概念是使用最佳值搜索類型取樣指標不斷地尋找最佳值，當取樣指標搜索到下一個位置與樣本之間的距離過近時，則利用減少誤差類型取樣指標來減少模型誤差，直到達到設定的迭代次數，或是誤差取樣指標的改善值過小，則演算停止。

最早的EGO演算法只針對無拘束條件的問題，後來主要有四種方法被提出來處理有拘束條件之問題的最佳化，分別是：機率法(Probability Method) [34]、懲罰函數(Penalty

Method) [20,33]、**違反期望值**(Expected Violation) [58]、**拘束型取樣指標**(Constrained ISC Method) [13]。

這些方法在參考文獻 [33]中有做比較，而我們將以拘束型取樣指標來作爲處理拘束條件的方法，然而在使用拘束型取樣指標方法時，必須注意到會有無可行解取樣指標的情形出現，在此狀況出現時，我們會以最大標準差指標(Maximum Variance)來取代本次迭代原本使用的指標，因此關於有拘束條件的演算策略，如圖5.3所示。

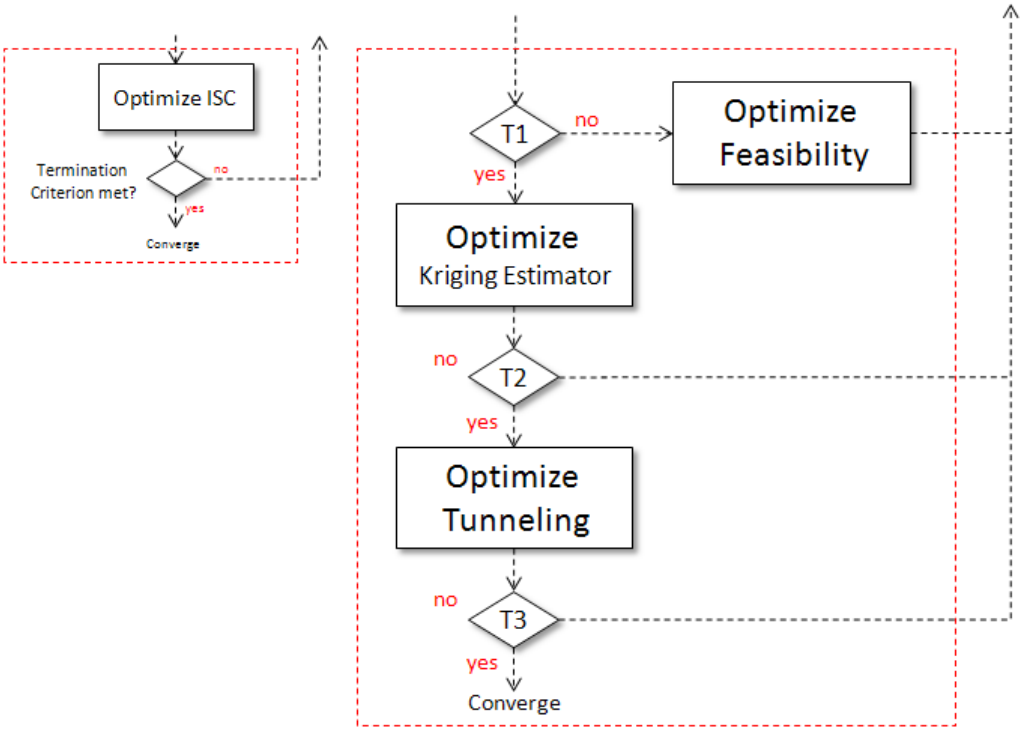


圖 5.3: 有拘束條件演算策略圖

將上圖5.3中的終止條件的內容列出，如下表5.2。

表 5.2: 有拘束條件演算策略之終止條件

編號	條件內容
1	檢查目前樣本中是否有可行解
2	改善目標函數指標與其他樣本過近
3	隧道指標數值小於使用者自訂數或是與其他樣本過近

此演算流程的主要概念是，在一開始檢查所有樣本中是否有可行解，如果沒有可行解

存在則使用可行解機率指標來搜索可行解，一旦存在可行解則使用最佳值搜索指標來尋找更好的解，當最佳值搜索指標無法找到更好的答案時，再使用隧道指標來尋找設計空間中其他可行解的存在。

5.2 最大可靠度問題 (Maximal Reliability Problem)

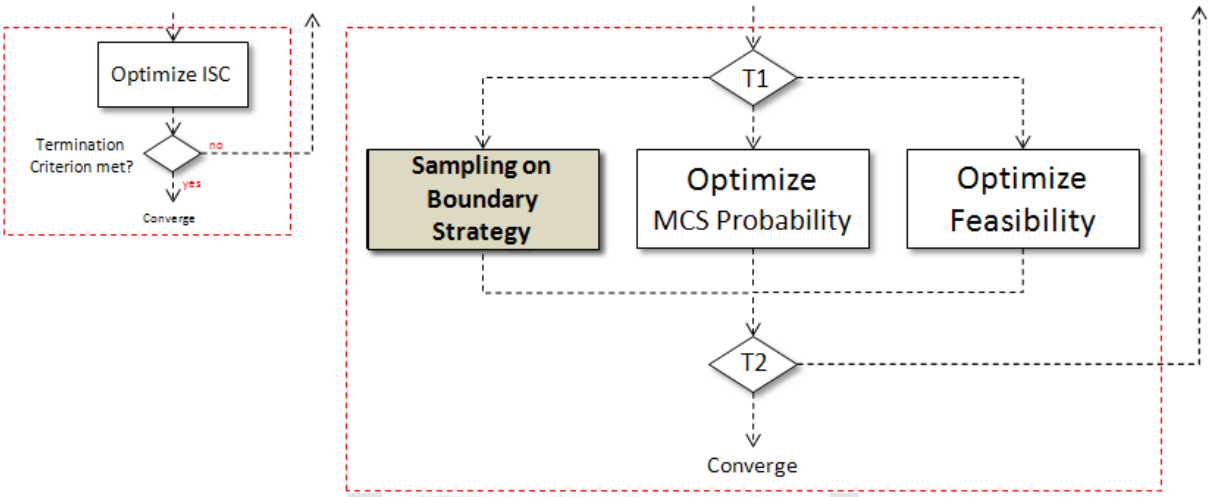


圖 5.4: 可靠度問題演算策略圖

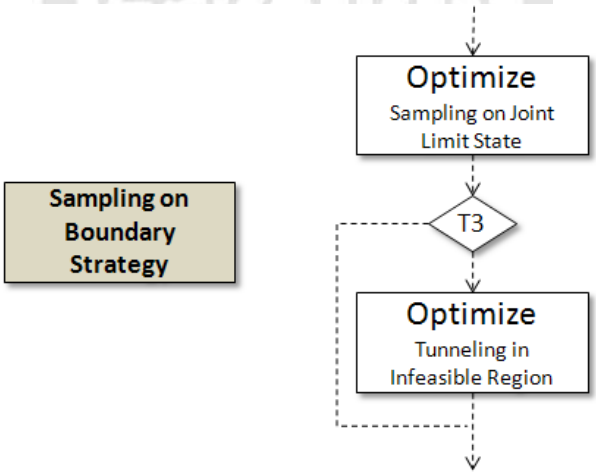


圖 5.5: 極限狀態取樣演算策略圖

如圖5.4，此爲考慮可靠度的問題之演算策略，其中關於極限狀態取樣的演算策略，如圖5.5，極限狀態取樣演算策略，顧名思義就是要試圖將問題的極限狀態進行修正，當極限狀態取樣指標取在與其它樣本過近的情況下，則利用隧道指標，在非可行解空間爲拘束條件的情況下，試圖尋找其他可行解空間，藉由這兩種取樣的方法，來進行對極限狀態的修正。將圖5.4與圖5.5中的終止條件的內容列出，如下表5.3。

表 5.3: 可靠度的問題演算策略之終止條件

編號	條件內容
1	檢查目前樣本中是否有可行解；目前樣本數目
2	得到的最大可靠度之位置與機率相差過小
3	極限狀態指標結果與其他樣本過近

此演算策略的主要概念是，一開始會先判斷是否存在可行解，若不存在則使用可行解機率指標來尋找可行解空間，一旦存在可行解則使用極限狀態取樣演算策略，試圖修正問題的極限狀態，然後在每十次迭代中，利用蒙地卡羅法對Kriging模型計算破壞機率，以最小化破壞機率爲目標函數的情況下來進行最佳化，當取樣指標位置與計算的破壞機率，與之前的結果相差過小，則演算停止。至此我們訂定了關於最大可靠度問題的演算策略，並在下一章範例說明中，實際應用這些演算策略。

第 六 章　範例 (Example)

6.1 凹形可行解空間問題 (Concave Feasible Region Problem)

參照第四章Baseball Function範例，此範例爲一個典型的凹形可行解空間問題，在此將此問題作爲最大可靠度問題的範例，將其重新撰寫如下列方程式(6.1)。

$$
\begin{aligned}
g_1 &= x_1^2 + x_2^2 - 0.8^2 \leq 0 \\
g_2 &= -x_1 + 2x_2^2) + 3x_1x_2 - 0.2 \leq 0 \\
g_3 &= x_1 - 2x_2^2 - 0.2 \leq 0
\end{aligned}
\tag{6.1}
$$

假設隨機變數$\mathbf{X}$，其分佈情形如方程式(6.2)。

$$
\mathbf{X} \backsim \left[N(\mu_{\mathrm{X}_1}, 0.2^2), N(\mu_{\mathrm{X}_2}, 0.15^2)\right] \tag{6.2}
$$

在使用DIRECT演算法與10^6個取樣點的蒙地卡羅法，直接對Baseball Function進行最大可靠度最佳化，其方程式(6.3)如下，其中$\max\left[\hat{\mathbf{P}}_{\mathrm{f}}(\boldsymbol{\mu}_{\mathrm{X}})\right]$代表著在設計點$\boldsymbol{\mu}_{\mathrm{X}}$，所有拘束條件所計算的破壞機率中的最大值，而我們的目標是要將其最小化。

$$
\min_{\boldsymbol{\mu}_{\mathrm{X}}} \quad \left\{\max\left[\hat{\mathbf{P}}_{\mathrm{f}}(\boldsymbol{\mu}_{\mathrm{X}})\right]\right\} \tag{6.3}
$$

最佳化後的結果，是在座標$(-0.06950, 0.40146)$有最小的破壞機率0.011197，將其位置、變數分佈與極限狀態繪製，如圖6.1，其中黑色實線爲極限狀態位置、$\bullet$爲每個取樣點位置，星號爲DIRECT結合蒙地卡羅法的最佳解位置。

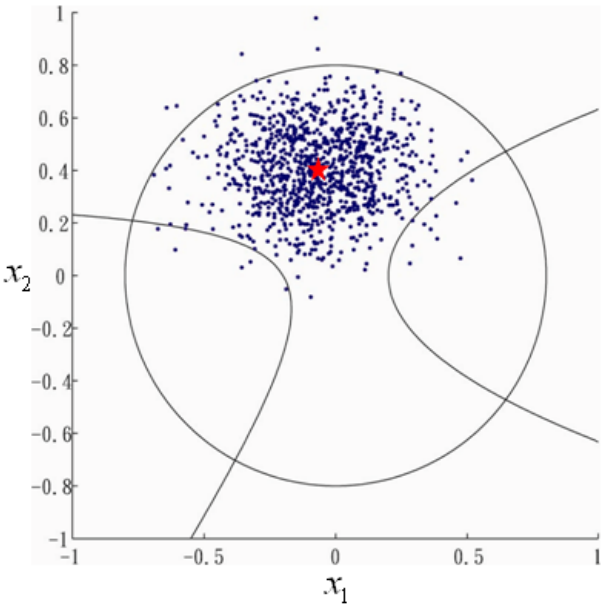


圖 6.1: DIRECT演算法之凹形可行解空間問題最大可靠度位置示意圖

利用第五章之最大可靠度演算策略，來對此凹形可行解空間問題進行最大可靠度最佳化，其結果如表6.1所示，表中的破壞機率是在不同樣本數下之最佳解位置，重新對原函數進行蒙地卡羅法10^6個樣本點測試結果，而此問題的驗證方法，是針對原函數使用DIRECT演算法加上蒙地卡羅法計算可靠度，每一次計算可靠度皆爲10^6個樣本點，最佳解爲DIRECT執行20207次函數計算次數的結果，因此實際上計算原函數的總次數爲2.0207×10^{10}次。 。

表 6.1: 凹形可行解空間問題最大可靠度最佳化結果表

迭代	樣本數	指標(LS/TU/MR)	最佳解位置	破壞機率
1	30	8/0/1	(-0.09968, 0.37586)	0.011446
2	40	9/0/1	(-0.09877, 0.37047)	0.012128
3	50	9/0/1	(-0.08505, 0.38043)	0.011080
4	60	9/0/1	(-0.09145, 0.38957)	0.010366
DIRECT			(-0.06313, 0.40420)	0.011039

初始樣本爲21個樣本點

上表中的指標(LS/TU/MR)代表著此迴圈中，使用了幾次不一樣的指標，其中各個代號所代表的內容如表6.2。

表 6.2: 指標代號表

指標代號	內容
LS	聯合極限狀態取樣指標
TU	隧道指標在非可行解空間
MR	蒙地卡羅法搜尋最大可靠度位置

由表6.1觀察到，在第60個樣本點所找到的最佳解，其破壞機率比DIRECT演算法還要低，歸咎其原因很有可能是因爲蒙地卡羅法的不確定性，事實上兩個答案只相差了大約700個樣本點，而另一點是DIRECT演算法不能重新計算已經計算過的位置，所以也無法確定DIRECT演算法的答案是否是最好的答案。

6.2 可靠度最佳化測試範例 (RBDO Test Problem)

參考文獻 [59]。此範例爲可靠度最佳化設計的一個典型的測試問題，我們簡稱爲RBDO範例，其拘束條件如方程式(6.4)，其圖形如圖6.2。

$$
\begin{aligned}
g_1 &= X_1^2 X_2 - 20 \leq 0 \\
g_2 &= 1 - \frac{(X_1 + X_2 - 5)^2}{30} - \frac{(X_1 + X_2 - 12)^2}{120} \leq 0 \\
g_3 &= X_1^2 + 8X_2 - 75 \leq 0
\end{aligned}
\tag{6.4}
$$

由於原問題的隨機變數分佈範圍過小，因此我們在此重新定義問題的隨機變數X，其分佈情形如方程式(6.5)。

$$
X_1 \backsim N(\mu_{\mathbf{X_1}}, 1^2),\ X_2 \backsim N(\mu_{\mathbf{X_2}}, 2^2) \tag{6.5}
$$

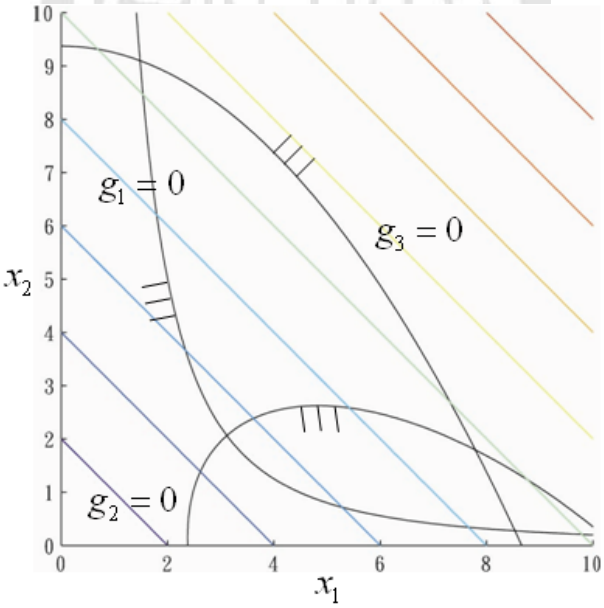
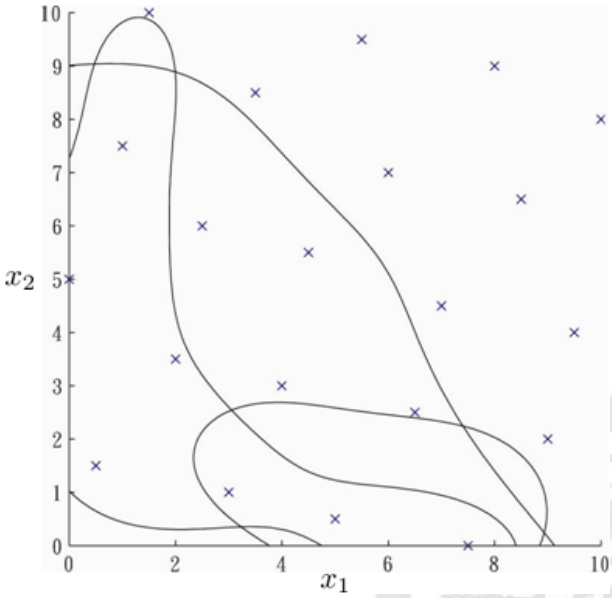
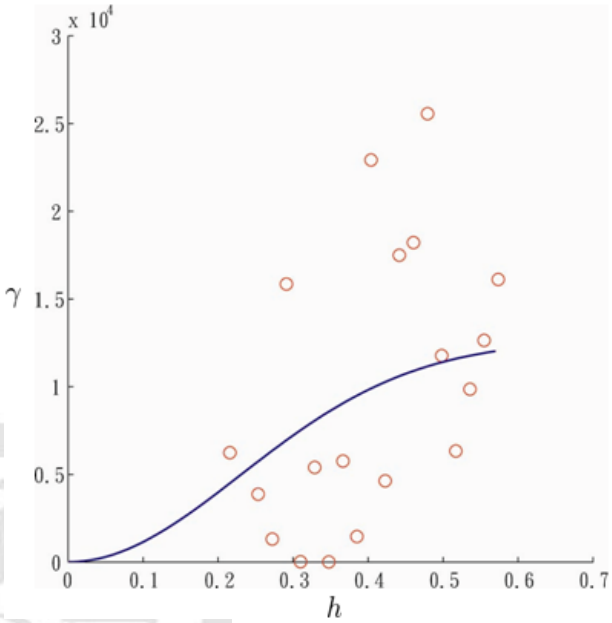


圖 6.2: 可靠度最佳化測試問題圖

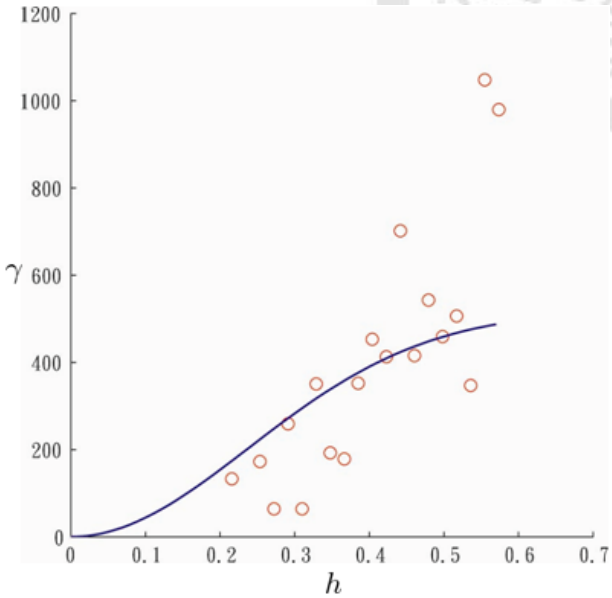
利用最大可靠度最佳化演算策略，以50個樣本數爲上限，進行對RBDO範例進行最大可靠度最佳化，其取樣位置與極限狀態的修正過程，以及各個拘束條件在變異圖擬合的過程，如圖6.3、圖6.4、圖6.5、圖6.6。



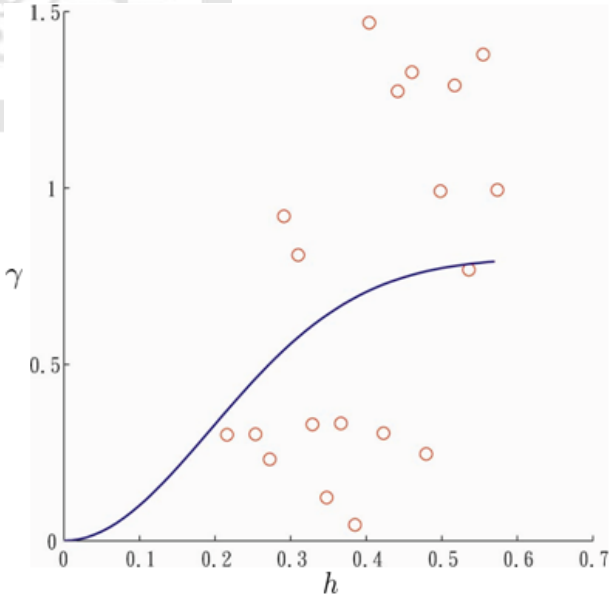
(a) 21個取樣極限狀態預測



(b) 21個取樣g_1之變異圖擬合

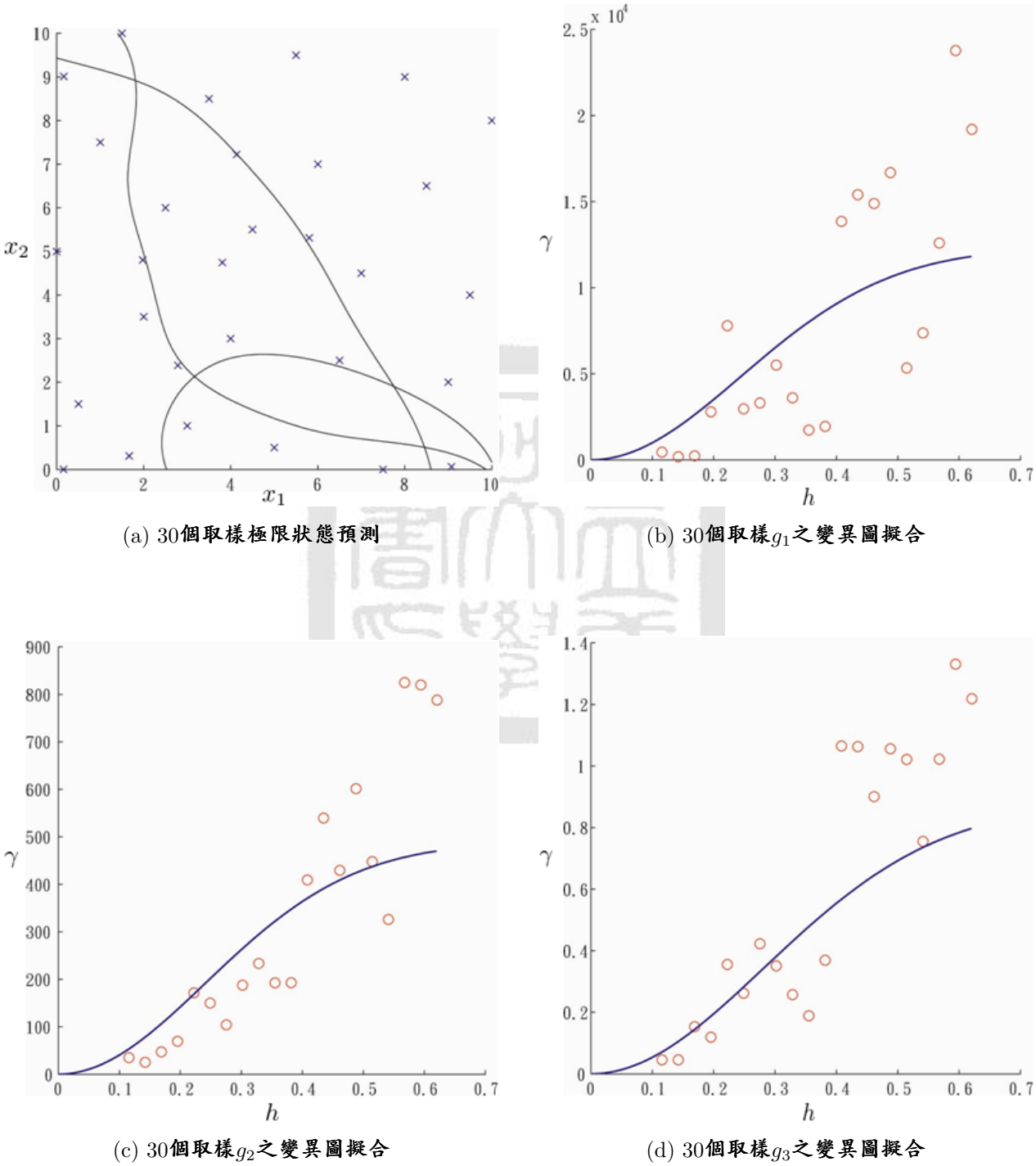


(c) 21個取樣g_2之變異數擬合



(d) 21個取樣g_3之變異數擬合

圖 6.3: RBDO**範例之**21**個樣本極限狀態圖**



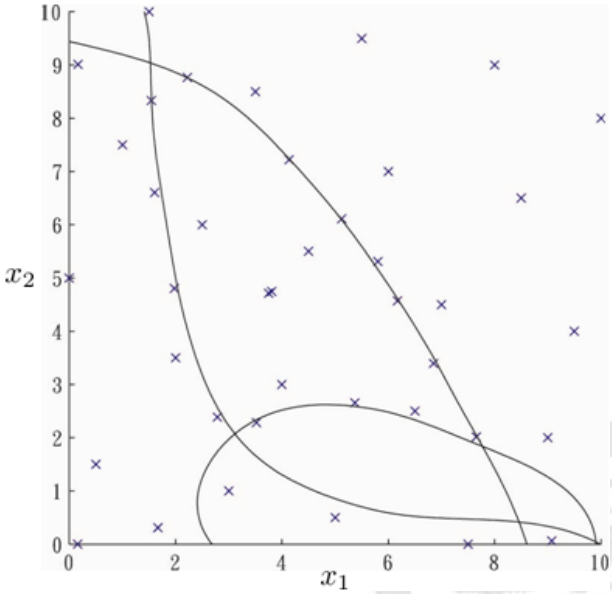
(a) 30個取樣極限狀態預測

(b) 30個取樣g_1之變異圖擬合

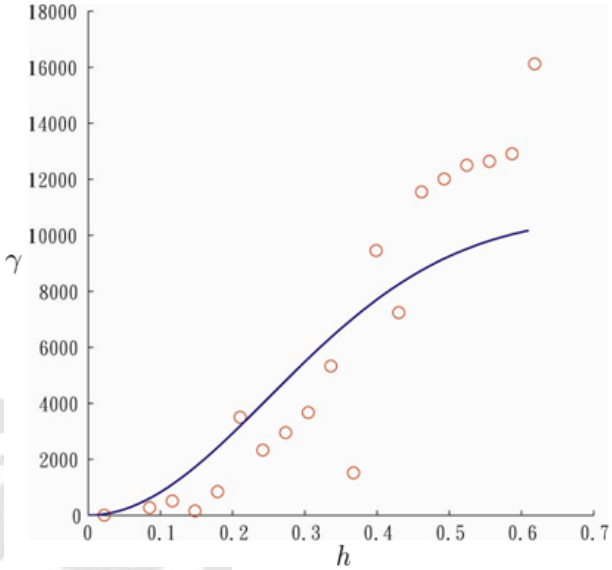
(c) 30個取樣g_2之變異圖擬合

(d) 30個取樣g_3之變異圖擬合

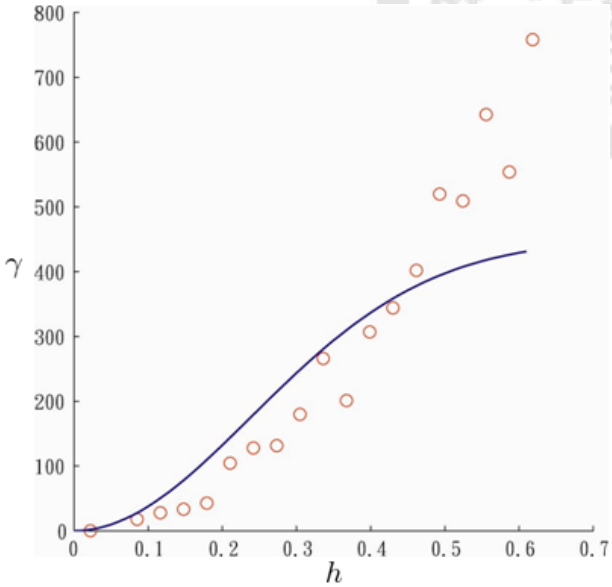
圖 6.4: RBDO**範例之**30**個樣本極限狀態圖**



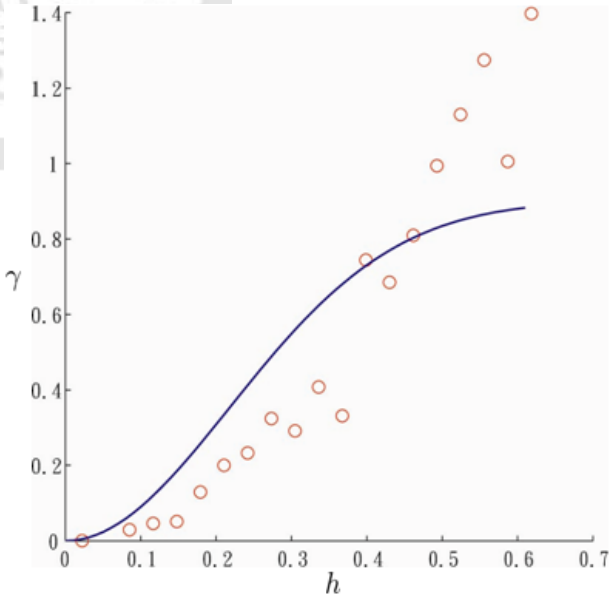
(a) 40個取樣極限狀態預測



(b) 40個取樣g_1之變異圖擬合

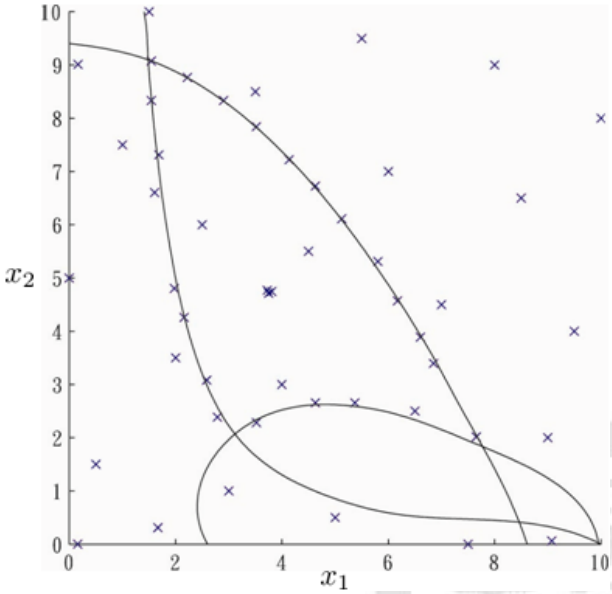


(c) 40個取樣g_2之變異圖擬合

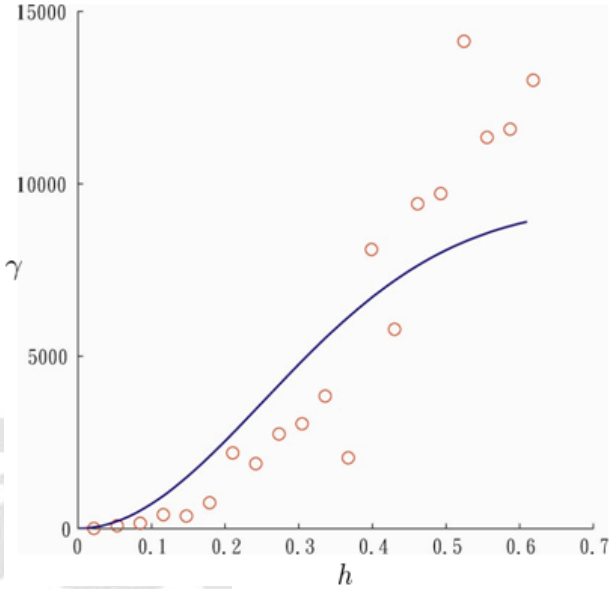


(d) 40個取樣g_3之變異圖擬合

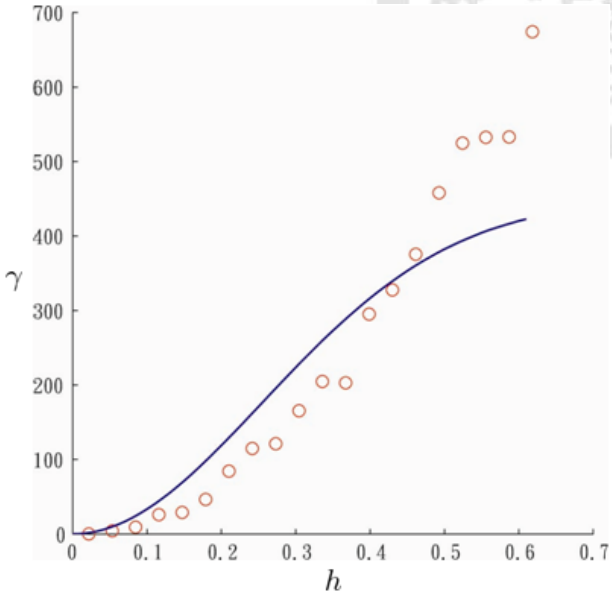
圖 6.5: RBDO**範例之**40**個樣本極限狀態圖**



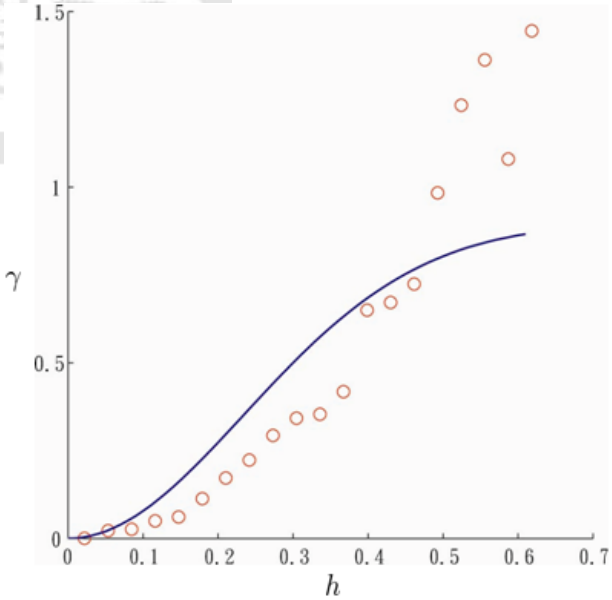
(a) 50個取樣極限狀態預測



(b) 50個取樣g_1之變異圖擬合



(c) 50個取樣g_2之變異圖擬合



(d) 50個取樣g_3之變異圖擬合

圖 6.6: RBDO**範例之**50**個樣本極限狀態圖**

我們在樣本數30、40、50時，分別對目前的Kriging模型進行最大可靠度最佳化，其結果如表6.3，表中指標的代號與表6.2，而此問題的驗證方法，是針對原函數使用DIRECT演算法加上蒙地卡羅法計算可靠度，每一次計算可靠度皆爲10^6個樣本點，最佳解爲DIRECT執行20187次函數計算次數的結果，因此實際上計算原函數的總次數爲2.0187×10^{10}次。

表 6.3: RBDO測試範例之最大可靠度最佳化結果表

迭代	樣本數	指標(LS/TU/MR)	最佳解位置	最大可靠度
1	30	8/0/1	(3.7473, 4.7317)	0.890635
2	40	9/0/1	(3.7512, 4.7529)	0.891101
3	50	9/0/1	(3.7379, 4.7302)	0.889680
	DIRECT		(3.7613, 4.7301)	0.892024

初始樣本爲21個樣本點

我們由實驗變異圖的結果來進行討論，題目的三個拘束條件在變數空間中，大致上都是屬於平滑且連續的函數，然而在一開始爲了得到整體函數的走向，我們使用Space-Filling的概念來作初始樣本的選擇，而Space-Filling的概念是要將樣本盡量分散在變數空間中，使得各個樣本之間的最小距離最大，因此對於21個初始樣本點的實驗變異圖來說，這種取樣方式會缺少了函數空間相關性中距離小的資訊，使得最小變異圖擬合法無法對拘束條件能有較完整的擬合。

6.3 調諧減震器問題 (Tuned Vibration Absorber Problem)

參考文獻 [60]，此範例將作爲極限狀態搜尋範例，並且將結果與文獻做比較。

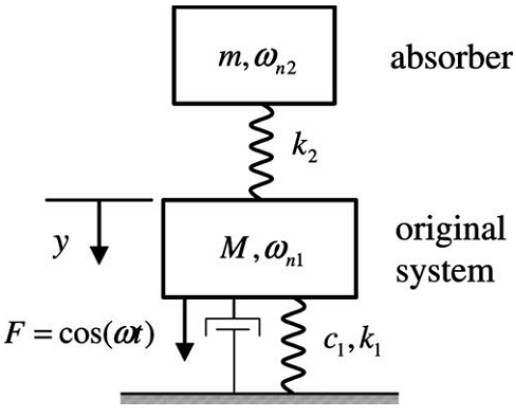


圖 6.7: Zou et al.所使用的調諧減震器示意圖

如圖6.7，此系統爲一個自由度的調諧阻尼器系統，包含了一個原系統與一個減震器(Vibration Absorber)，加裝了減震器的主要目的是要讓整體系統的的振動幅度減小，而此系統的振動幅度爲y，可以用下列方程式(6.6)表示，其中ω_{n_1}與ω_{n_2}分別代表原系統與減震器的自然頻率(Nature Frequency)，ω爲減震器的激發頻率(Excitation Frequency)，而R代表減震器與原系統的質量比，ς代表原系統的阻尼比，此題目限定振動幅度y不可超過28，因此拘束條件可寫成方程式(6.7)，其中問題的隨機變數分佈皆爲高斯分佈。

$$y = \frac{1 - \left(\frac{1}{X_2}\right)^2}{\sqrt{\left[1 - R\left(\frac{1}{X_1}\right)^2 - \left(\frac{1}{X_1}\right) - \left(\frac{1}{X_2}\right)^2 + \frac{1}{X_1^2 X_2^2}\right]^2 + 4\varsigma^2\left[\frac{1}{X_1} - \frac{1}{X_1 X_2^2}\right]^2}} \tag{6.6}$$

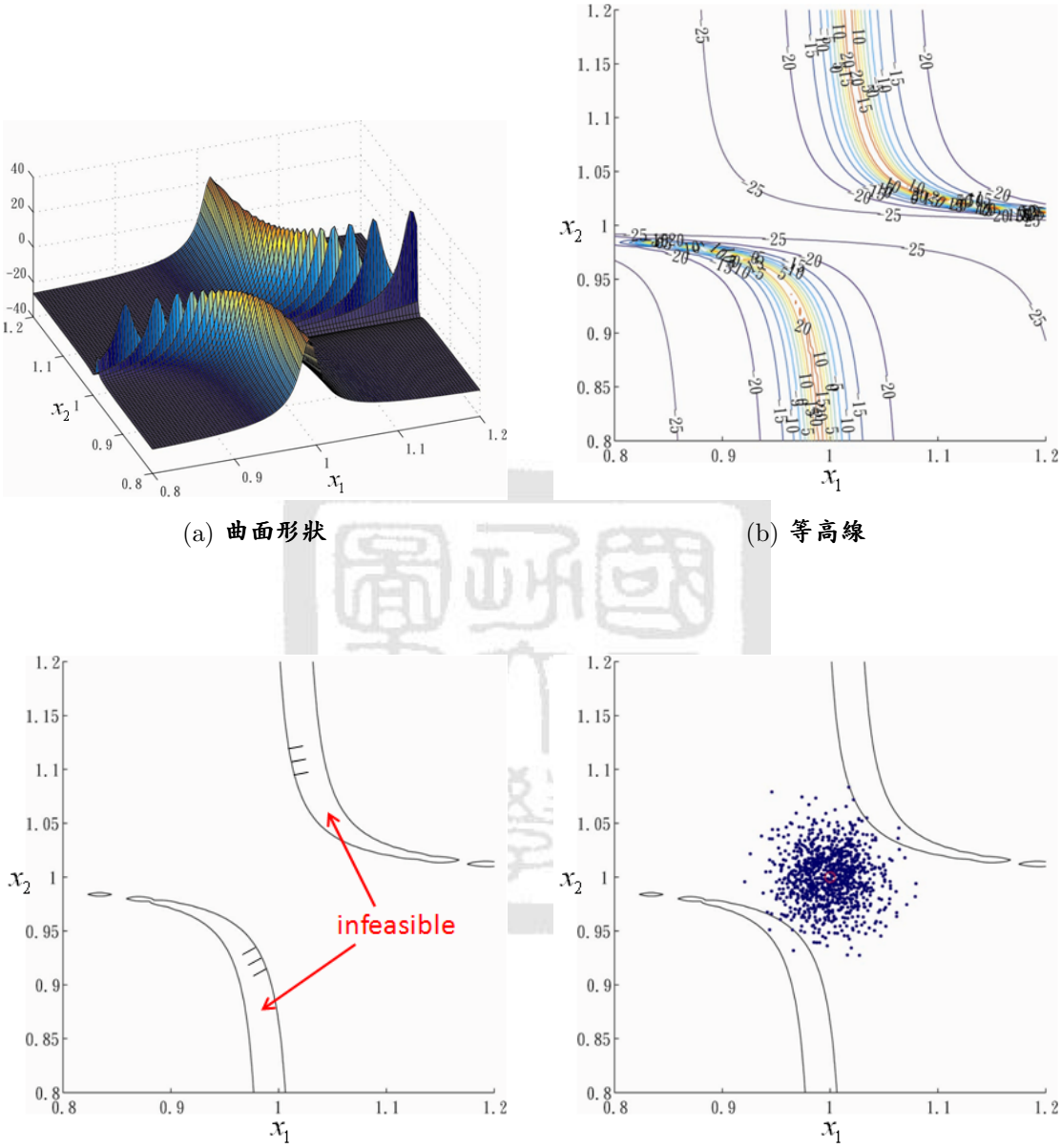
$$X_1 = \frac{\omega_{\mathrm{n}1}}{\omega} \backsim N(1, 0.025), \quad X_2 = \frac{\omega_{\mathrm{n}2}}{\omega} \backsim N(1, 0.025)$$

$$R = 0.01, \quad \varsigma = 0.01$$

$$g(X_1, X_2) = y - 28 \leq 0 \tag{6.7}$$

將其拘束條件的形狀、等高線、極限狀態以及問題的隨機變數分佈範圍示意繪製，如

圖6.8，其中圖6.8-(d)是隨機變數分佈範圍在取樣1000點的情況下的示意圖。

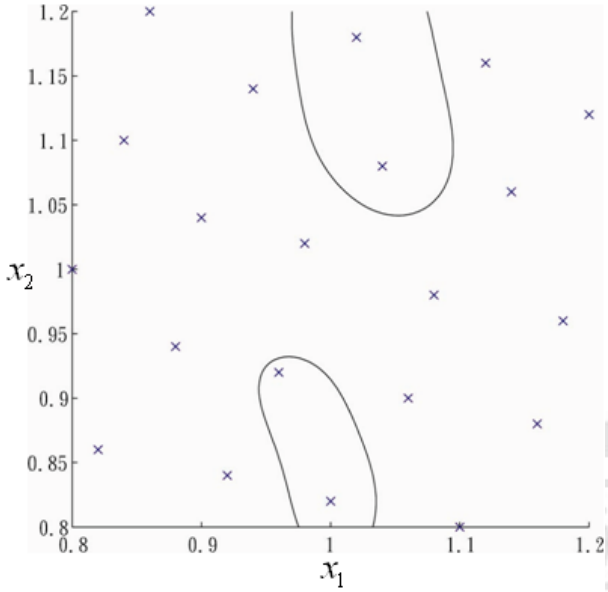


(a) 曲面形狀　(b) 等高線

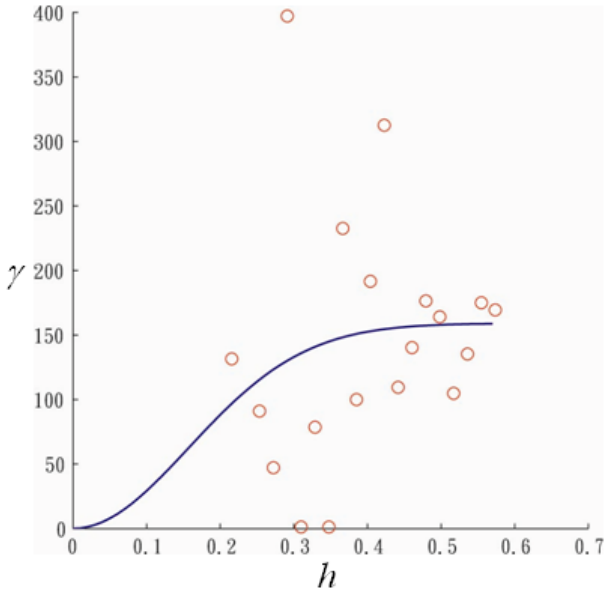
(c) 極限狀態　(d) 隨機變數分佈

圖 6.8: 減震器工程範例問題示意圖

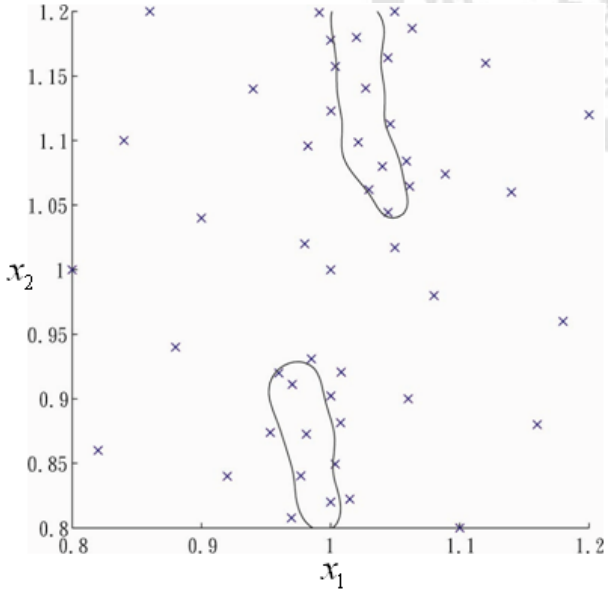
利用極限狀態取樣指標，試圖將此問題的極限狀態進行修正，其修正的過程如圖6.9、圖6.10、圖6.11、圖6.12，其中左圖爲取樣指標運算結果，右圖爲在目前樣本極限狀態的預測位置。



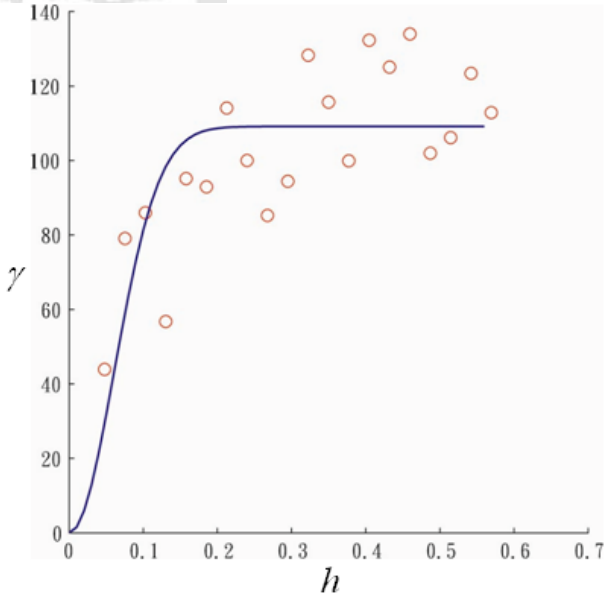
(a) 21個樣本變異圖擬合



(b) 21個樣本極限狀態預測

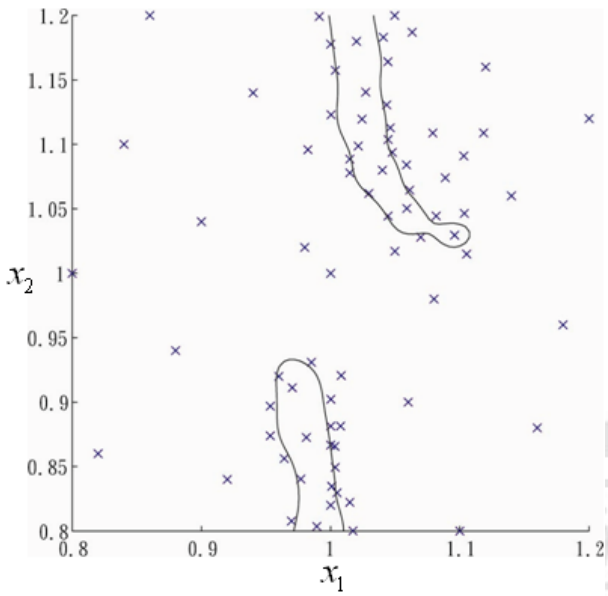


(c) 50個樣本變異圖擬合

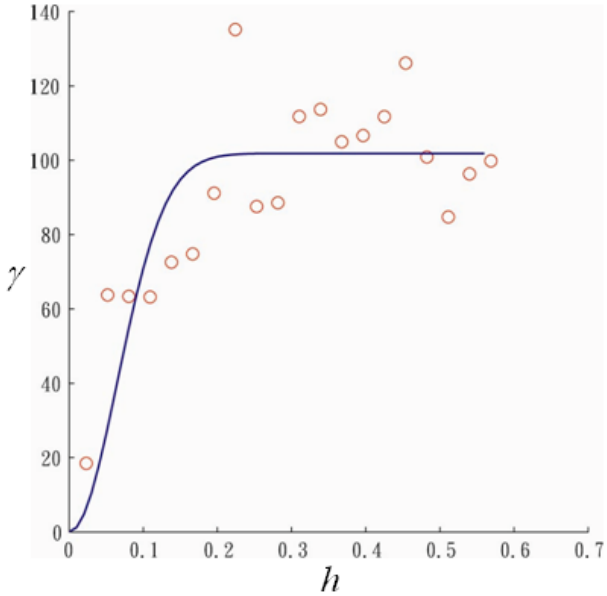


(d) 50個樣本極限狀態預測

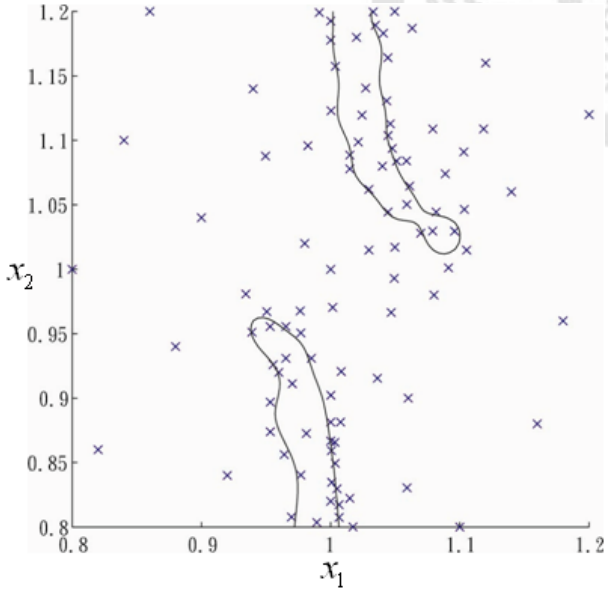
圖 6.9: 減震器工程範例極限狀態修正過程圖(21、50)



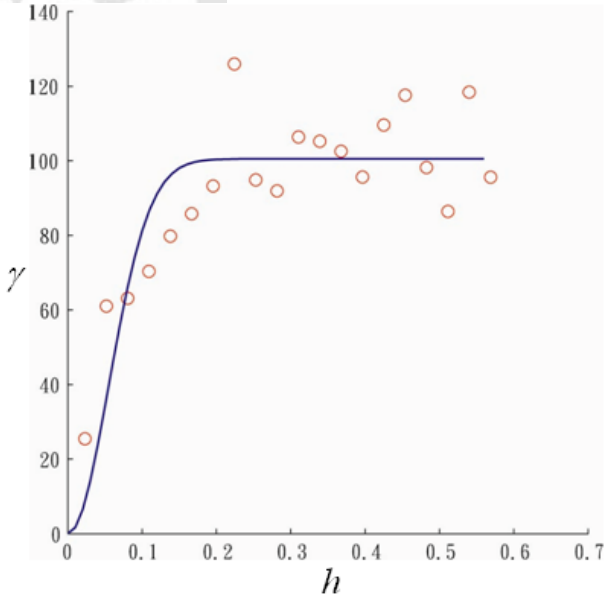
(a) 75個樣本變異圖擬合



(b) 75個樣本極限狀態預測

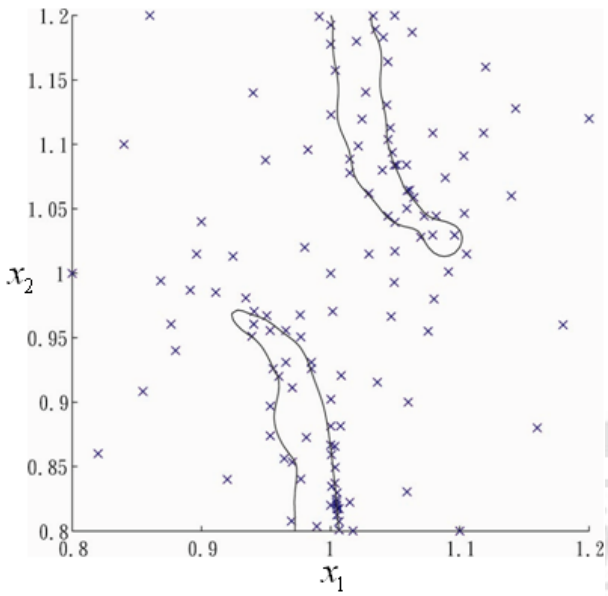


(c) 100個樣本變異圖擬合

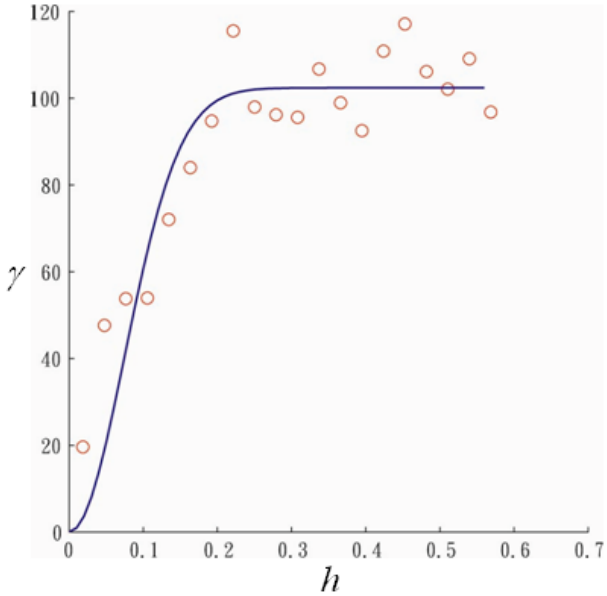


(d) 100個樣本極限狀態預測

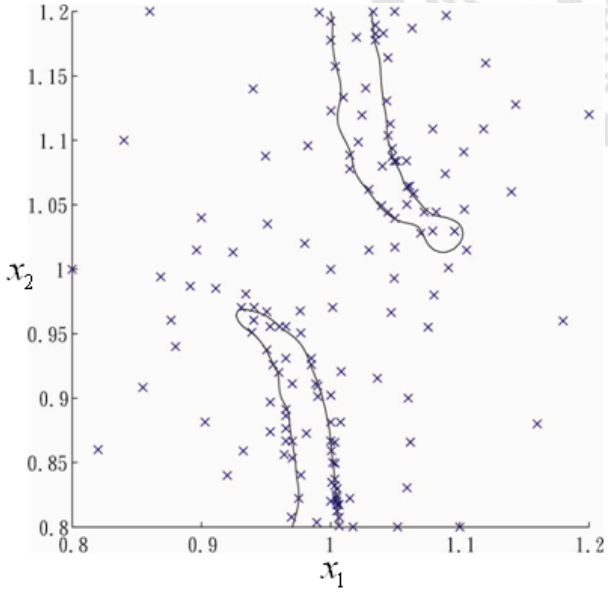
圖 6.10: 減震器工程範例極限狀態修正過程圖(75、100)



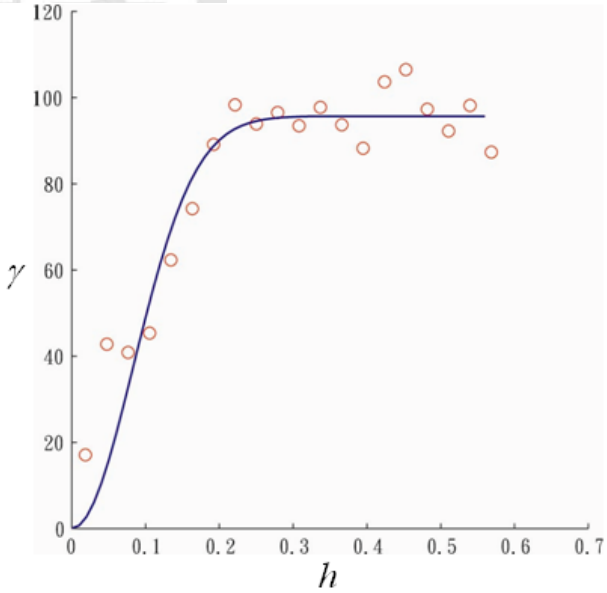
(a) 125個樣本變異圖擬合



(b) 125個樣本極限狀態預測

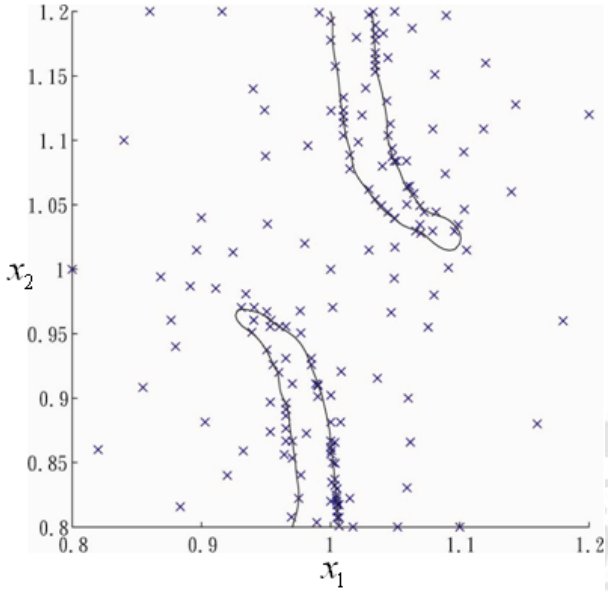


(c) 150個樣本變異圖擬合

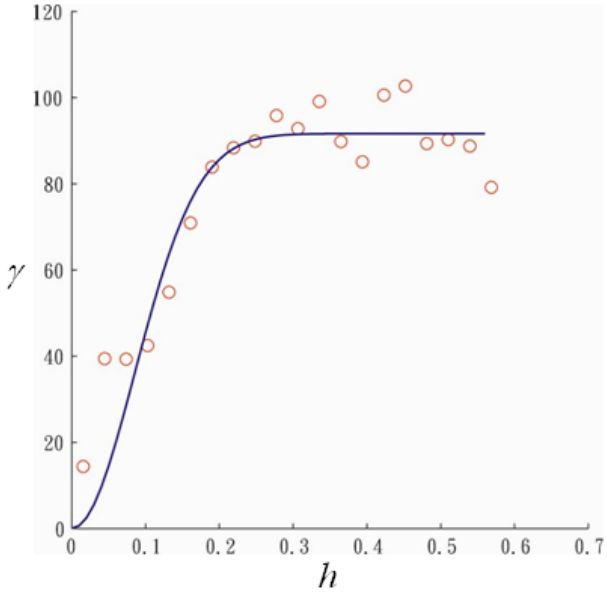


(d) 150個樣本極限狀態預測

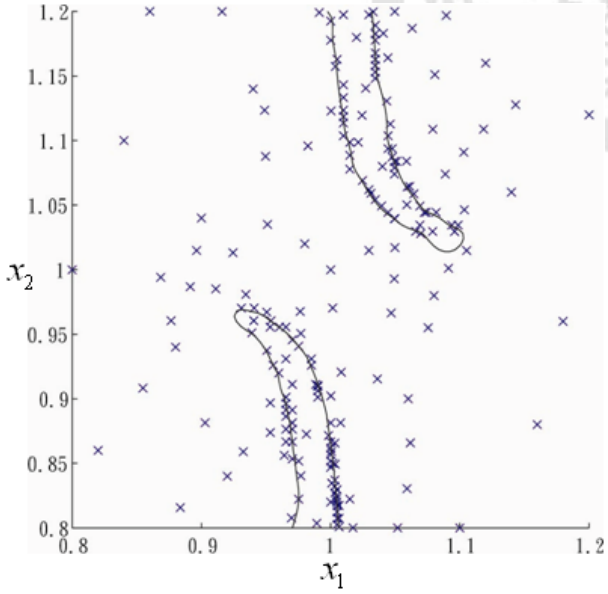
圖 6.11: 減震器工程範例極限狀態修正過程圖(125、150)



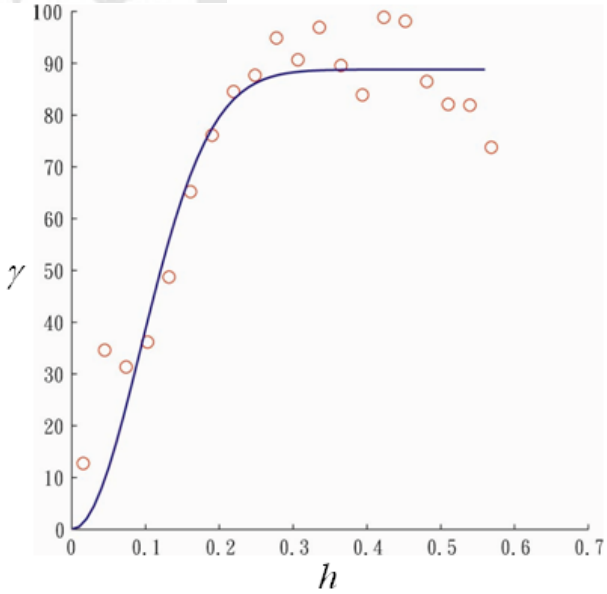
(a) 175個樣本變異圖擬合



(b) 175個樣本極限狀態預測



(c) 200個樣本變異圖擬合



(d) 200個樣本極限狀態預測

圖 6.12: **減震器工程範例極限狀態修正過程圖**(175、200)

依照問題之隨機變數分佈，利用蒙地卡羅法在計算原數學函數與Kriging模型同時取樣10^5個樣本點來計算破壞機率，比較文獻 [60]與我們的結果，如圖6.13、表6.4。

Table 1 IRS-MC results for vibration absorber problem

Limit-state function: $g=28.0-y(\beta_1,\beta_2)$
Distribution of random variables: $\beta_1 \sim N(1,0.025)$, $\beta_2 \sim N(1,0.025)$

Probability of failure from Monte Carlo simulation: 0.0097
(30,000 sample points)

	P_f	$\frac{P_f}{P_{f,\mathrm{MCS}}}$	No. of function evaluations
FORM	No convergence	—	—
Multimodal AIS (mean value as starting point)	0.00407	0.38	840
IRS-based Monte Carlo	0.00963	0.99	206[a]

[a]206=25(initial OSLH)+44(additional OSLH)+25(additional random)+112(final MCS).

圖 6.13: Zou et al.之減震器工程範例的結果圖

表 6.4: 減震器工程範例破壞機率計算結果表

樣本數目	指標(LS/TU)	$\hat{P}_{\mathrm{f}}$	$\frac{\hat{P}_{\mathrm{f}}}{\hat{P}_{\mathrm{f,MCS}}}$
21	-	0.00972	0.95
50	26/3	0.00457	0.45
75	18/7	0.00556	0.54
100	13/16	0.01017	0.99
125	10/19	0.01101	1.08
150	3/26	0.01097	1.07
175	5/24	0.01116	1.09
200	4/25	0.01075	1.05

$\hat{P}_{\mathrm{f,MCS}} = 0.01023$

初始取樣爲21點

由測試的結果來看，在100個樣本數目以上之後，大致上能得到與文獻上相去不遠的結果，但是還是可以由極限狀態修正過程圖看出，得到的極限狀態預測位置仍然與原函數有所差異，特別是在細小的極限狀態區域，因此在此設計點所計算的破壞機率能與原函數的破壞機率相同，並不代表在整個空間中的任意位置的破壞機率也會相同。然而必須注意到在文獻中，如表6.4，對原函數進行蒙地卡羅測試時，是取3×10^4個樣本點來計算，而我們的測試是以10^5個樣本點來計算，如果以10^6個樣本點對原函數進行蒙地卡羅計算的結果，其破壞機率爲0.010246，也就是說在文獻 [60]中，對於原函數的破壞機率計算，似乎是在蒙地卡羅的取樣點數不足，造成對破壞機率的低估。

另外由拘束條件的實驗變異圖結果觀察，由於演算策略是試圖在極限狀態上取樣本，因此理論上初始取樣之後所得到的樣本，函數值彼此之間不會有太大差異，然而在實驗變異圖的圖中卻可以發現到，擬合後的空間相關性影響範圍相當小，這代表著函數本生具有不連續或是在某處有峰值的產生，此種現象與圖6.8的說明大致相同，這也代表著我們可以藉由變異圖的分析，來推估函數在設計空間中所具有的特性，因此要對此種具有峰值的函數進行推估，必須要在峰值附近進行取樣才能得到較完整的資訊。

第 七 章　結論 (Conclusions)

本論文將原本只針對定量變數最佳化的EGO演算法，延伸至可靠度最佳化的問題，並對各種不同種類的問題使用不同的演算策略，提出一種解決具有黑盒子函數的可靠度最佳化問題的方法，而此黑盒子函數指的是由電腦模擬或是實際實驗等耗費時間才能得到數據的方法。以下將整理本論文之研究貢獻，並對本論文之研究內容提出一些探討，以供未來研究發展之參考。

7.1　研究貢獻 (Contribution)

本論文之研究貢獻，主要可以分爲下列五點：

1. 提出一種在使用最大似然推估法擬合Kriging模型的問題，此問題並不是因爲似然性函數或是問題本身的特性，而是一種在數值運算上會產生的現象，但是此現象在過去的文獻中並未被探討。
2. 制定一個自動化擬合變異圖的流程來建構Kriging模型，此變異圖擬合的概念源自於地理統計學。
3. 將過去文獻中所使用的取樣指標整理，依照其目的加以分類，並將這些取樣指標加入程式中，以供使用者選擇上的方便。
4. 提出一種新的取樣指標，而此指標屬於特殊目的類型的指標，能使EGO演算法在拘束條件上的極限狀態進行取樣，以修正Kriging模型與原函數在極限狀態上的誤差。
5. 制定一種新的EGO演算法演算流程，將原本只針對定率變數最佳化的演算法，延伸至可靠度最佳化的問題。

7.2 建議 (Suggestion)

在第二章所提到，EGO演算法的主要內容在於Kriging模型與取樣指標，因此瞭解這兩個項目的內容是有其必要性的；過去在機械領域使用Kriging模型的文獻大多延續於DACE所使用的內容，然而眞正Kriging模型理論背景是源自於地理統計領域，因此對於地理統計領域的不瞭解，將有可能導致我們對Kriging模型的誤用。

在第三章Kriging模型的基本概念中曾經提到過，本篇論文中都將Kriging模型假設爲高斯歷程，然而事實上Kriging模型並沒有假設爲某種形式的序率歷程，如此假設的原因是因爲高斯歷程的對稱性與計算方便，可以使得我們能夠很快速的計算取樣指標。而在問題與討論中還提到，DACE使用的最大似然推估法，將所有樣本資料純粹以數據的方式來對Kriging模型進行擬合，缺乏空間相關性在本質合理性與計算可行性上的考量，然而這樣的作法是否合宜我們並不瞭解，因爲對於一個設計問題來說，變數與變數之間對於函數的關係，並不像地理統計領域中地理位置方向與函數的關係那麼明顯，舉例來說，參考文獻 [61]中設計一個煞車組件，如圖7.1，如果控制因子爲煞車墊片(Pad)及轉輪(Rotor)的參數，那麼我們很難想像墊片形狀與轉輪材料對於煞車系統的機械效率，在變數空間中有何種相關性。

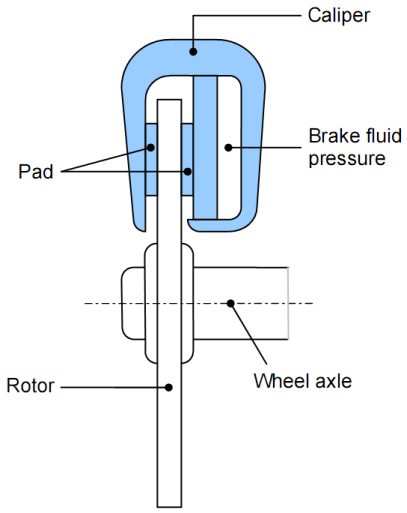


圖 7.1: Lee et al.所使用的煞車組件示意圖

如果將最大似然推估法所使用的空間相關函數對比於地理統計學中的概念，所謂的 *product correlation rule* 很類似於非等向性(Anisotropy)變異圖的概念，這似乎解決了我

們難以判斷樣本在變數空間中的相關性問題，然而最大似然推估法在使用上又存在著一些不可避免的現象，如同第三章討論到，當樣本數一多時，計算似然函數中的樣本空間相關係數矩陣的行列式值 $\det(\mathbf{R}_\alpha)$ ，將會超出電腦所能儲存的資料型態上限，這會造成似然函數在使用上的困難，這也是本論文在進行Kriging模型的擬合時，使用最小變異圖平方誤差法的原因。

另外一點是，在我們所提出來的極限狀態取樣指標，其本身對於一些最佳化演算法來說並不是一個很好的目標函數，因爲此指標就如同EI指標、Kushner指標、Feasibility指標等，都具有一個不好的性質，就是具有函數值爲零的大區域以及一些峰值，而不像Lower Confidence Bounding Function指標、Kriging Estimator指標等，具有完整空間資訊的指標，這會造成演算法在進行搜尋上的困難，如果能夠改善極限狀態取樣指標的此性質，將有助於改善在極限狀態上的搜尋效率。

事實上我們對於Kriging模型的使用，還是只局限於地理統計學中的一小部份，然而在Kriging模型中，還有IRFk-Kriging、Indicator Kriging、Disjunctive Kriging、Lognormal Kriging、Constrained Kriging等其他種模式的發展，可參考 [36]。因此，我們似乎還可以從地理統計領域中得到一些方法來針對我們的問題進行應用。例如：針對Cammelback函數，以拘束條件作爲門檻(Threshold)來繪製Indicator Kriging，如圖7.2，圖中實線爲問題的極限狀態，x爲取樣位置，等高線可視爲Indicator Kriging預測函數值大於零的機率爲多少。

對於可靠度問題來說，破壞機率的計算往往需要有設計值與極限狀態的距離、或是在設計值周遭的隨機變數分佈區域的資訊，此種需求與地理統計學中Kriging模型的區域化變數似乎能夠類比，而在本論文中將樣本取樣在極限狀態上，以修正Kriging模型的概念，與地理統計學中的Preferential Sampling的概念類似，因此在後續的研究裡，可以考慮往地理統計學中的概念來得到更多的應用。

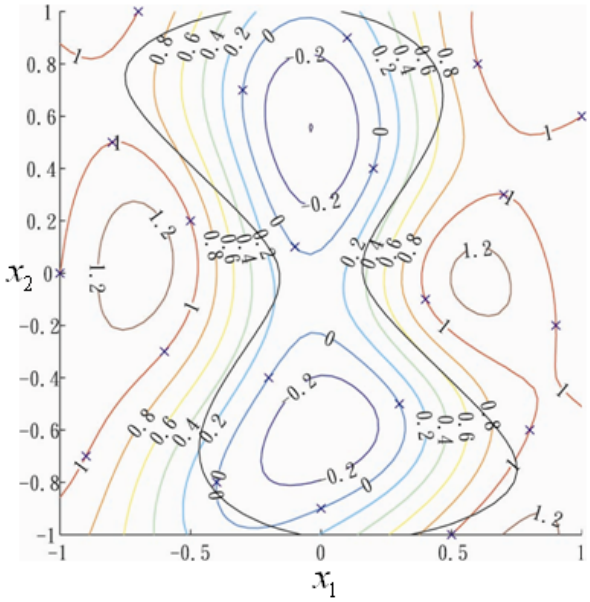


圖 7.2: Camelback函數的Indicator Kriging圖

參 考 文 獻

[1] J. Tu, K. K. Choi, and Y. H. Park, "A new study on reliability-based design optimization," *Journal of Mechanical Design*, vol. 121, no. 4, pp. 557–564, 1999.

[2] L. C. Dixon and G. P. Szegö, *Towards global optimization 2.* North- Holland, 1978.

[3] A. M. Hasofer and N. C. Lind, "Exact and invariant second-moment code format," *Journal of the Engineering Mechanics Division*, vol. 100, pp. 111–121, January/February 1974.

[4] P. Liu and A. Kiureghian, "Optimization algorithm for structure reliability," *Structural Safety*, vol. 9, pp. 161–177, 1991.

[5] J. Tu and K. Choi, "A new study on reliability-based design optimization," *Journal of Mechanical Design*, vol. 121, no. 4, pp. 557–564, 1999.

[6] J. Tu, K. K. Choi, and Y. H. Park, "Design potential method for robust system parameter design," *AIAA Journal*, vol. 39, no. 4, pp. 667–677, 2001.

[7] A. Haldar and S. Mahadevan, *Probability, Reliability, and Statistical Method in Engineering Design.* John Wiley, Sons, Inc, 2000.

[8] Y.-G. Zhao and T. Ono, "New point estimates for probability moments," *Journal of Engineering Mechanics*, vol. 126, pp. 433–436, April 2000.

[9] A. Harbitz, "An efficient sampling method for probability of failure calculation," *Structural Safety*, vol. 3, no. 2, pp. 109–115, 1986.

[10] D. R. Jones, M. Schonlau, and W. J. Welch, "Efficient global optimization of expensive black-box functions," *Journal of Global Optimization*, vol. 13, pp. 455–492, 1998.

[11] D. R. Jones, "A taxonomy of global optimization methods based on response surfaces," *Journal of Global Optimization*, vol. 21, pp. 345–383, 2001.

[12] H. M. Gutmann, "A radial basis function method for global optimization," *Journal of Global Optimization*, vol. 19, pp. 201–227, 2001.

[13] M. J. Sasena, P. Y. Papalambros, and P. Goovaerts, "The use of surrogate modeling algorithms to exploit disparities in function computation time within simulation-based optimization," in *In The Fourth World Congress of Structural and Multidisciplinary Optimization*, 2001.

[14] J. Knowles, "ParEGO : A hybrid algorithm with on-line landscape approximation for expensive multiobjective optimization problems," *IEEE Transactions on Evolutionary Computation*, vol. 10, pp. 50–66, 2006.

[15] M. Locatelli, "Bayesian algorithms for one-dimensional global optimization," *Journal of Global Optimization*, vol. 10, no. 1, pp. 57–76, 1997.

[16] H. J. Kushner, "A new method of locating the maximum of an arbitrary multipeak curve in the presence of noise," *Journal of Basic Engineering*, vol. 86, pp. 97–106, March 1964.

[17] B. Betrò, "Bayesian methods in global optimization," *Journal of Global Optimization*, vol. 1, no. 1, pp. 1–14, 1991.

[18] M. J. Sasena, P. Y. Papalambros, and P. Goovaerts, "Global optimization of problems with disconnected feasible regions via surrogate modeling," in *9th AIAA/ISSMO Sympoium on Multidisciplinary Analysis and Optimization*, AAIA 2002-5573, (Atlanta, Georgia), September 2002.

[19] S. Gomez and A. V. Levy, "The tunneling method for solving the constrained global optimization problem with several non-connected feasible regions," *Lecture Notes in Mathematics*, vol. 909, pp. 34–47, 1982.

[20] M. J. Sasena, P. Y. Papalambros, and P. Goovaerts, "Metamodeling sampling criteria in global optimization framework," in *8th AIAA/USA/NASA/ISSMO Symposium on Multidisciplinary Analysis and Optimization*, (Long Beach, CA), American Institude of Aeronautics & Astronautics, 6-8 September 2000.

[21] F. H. Branin, "Widely convergent method for finding multiple solutions of simultaneous nonlinear equations," *IBM Journal of Research and Development*, vol. 16, pp. 504–522, 1972.

[22] D. G. Krige, "A statistical approach to some mine valuation and allied problems on the Witwatersrand," Master's thesis, University of Witwatersrand, 1951.

[23] F. P. Agterberg, "Georges Matheron : Founder of spatial statistics," *Earth Sciences History*, vol. 23, no. 2, pp. 205–334, 2007.

[24] N. Cressie, "The origin of Kriging," *Mathematical Geology*, vol. 22, no. 3, pp. 239–252, 1990.

[25] J. Sacks, S. B. Schiller, and W. J. Welch, "Design for computer experiments," *Technometrics*, vol. 31, pp. 41–47, 1989.

[26] J. Sacks, W. J. Welch, W. J. Mitchell, and H. P. Wynn, "Design and analysis of computer experiments," *Statistical Science*, vol. 4, no. 4, pp. 409–435, 1989.

[27] A. J. Booker, J. E. Dennis, P. D. Frank, D. B. Serafini, and V. Torczon, "Optimization using surrogate objectives on a helicopter test example," tech. rep., Applied Research Technology, 1997.

[28] K. Park, P. K. Oh, and H. Lim, "The application of the CFD and Kriging method to an optimization of heat sink," *International Journal of Heat and Mass Transfer*, no. 49, pp. 3439–3447, 2006.

[29] S. Sakata, F. Ashida, and M. Zako, "Eigenfrequency optimization of stiffened plate using Kriging estimation," *Computational Mechanics*, no. 31, pp. 409–418, 2003.

[30] C.-H. Liu and K.-Y. Chan, "Reliability-based design of a pressure tank under random and stochastic environments," ASME International Design Engineering Technical Conference, Auguest,3-6 2008.

[31] L. Etman, "The method of Sacks, Design and analysis of computer experiments," tech. rep., Eindhoven University of Technology, 1994.

[32] P. K. Kitanidis, *Introduction to Geostatistics*. Cambridge University Press, 1997.

[33] M. J. Sasena, *Flexibility and Efficiency Enhancements for Constrained Global Design Optimization with Kriging Approximation*. PhD thesis, University fo Michigan, 2002.

[34] M. Schonlau, *Computer experiments and global optimization.* PhD thesis, University of Waterloo, 1997.

[35] J. D. Martin and T. W. Simpson, "Use of Kriging models to approximate deterministic computer models," *AIAA Journal*, vol. 43, no. 4, pp. 853–863, 2005.

[36] J. P. Chiles and P. Delfiner, *Geostatistics : Modeling Spatial Uncertainty.* Wiley-Interscience, 1999.

[37] J. L. Devore, *Probability and Statistics.* Thomson Brooks/cole, 2008.

[38] C. Currin, T. Mitchell, M. Marris, and D. Ylvisaker, "A bayesian approach to the design and analysis of computer experiments," tech. rep., Oak Ridge National Laboratory, 1989.

[39] K. V. Mardia and R. J. Marshall, "Maximum likelihood estimation of models for residual covariance in spatial regression," *Biometrika*, vol. 71, no. 1, pp. 135–146, 1984.

[40] D. G. Cook and S. J. Pocock, "Multiple regression in geographical mortality studies, with allowance for spatially correlated errors," *Biometrika*, vol. 39, pp. 361–371, 1983.

[41] T. H. O'Donnell, H. L. Southall, and L. B. Kaanta, "Efficient global optimization of a limited parameter antenna design," vol. 6964, Evolutionary and Bio-Inspired Computation: Theory and Applications II, March, 17.

[42] A. Todoroki and M. Sekishiro, "Modified efficient global optimization for a hat-stiffened composite panel with buckling constraint," *AIAA Journal*, vol. 46, no. 9, pp. 2257–2264, 2008.

[43] D. Huang, T. T. Allen, W. I. Notz, and N. Zeng, "Global optimization of stochastic black-box systems via sequential Kriging meta-models," *Journal of Global Optimization*, vol. 34, pp. 441–466, 2006.

[44] M. J. Sasena, P. Papalambros, and P. Goovaerts, "Exploration of metamodeling sampling criteria for constrained global optimization," *Engineering Optimization*, vol. 34, pp. 263–278, 2002.

[45] M. J. Sasena, M. Parkinson, M. P. Reed, P. Papalambros, and P. Goovaerts, "Improving an ergonomics testing procedure via approximation-based adaptive experimental design," *Journal of Mechanical Design*, vol. 127, pp. 1006–1013, 2005.

[46] C. Currin, T. Mitchell, M. Marris, and D. Ylvisaker, "Bayesian prediction of deterministic functions, with application to the design and analysis of computer experiments," *Journal of the American Statistical Association*, vol. 86, pp. 953–963, December 1991.

[47] J. J. Warenw and B. D. Ripley, "Problems with likelihood esitmation of covariance functions of spatial gaussian processes," *Biometrika*, vol. 74, no. 3, pp. 640–642, 1987.

[48] G. J. Davis and M. D. Morris, "Six facotrs which affect the condition number of matrices associated with Kriging," *Mathematical Geology*, vol. 29, no. 5, pp. 669–683, 1997.

[49] S. E. Gano, J. E. Renaud, J. D. Martin, and T. W. Simpson, "Update strategies for Kriging models used in variable fidelity optimization," *Structure Multidisciplinary Optimization*, vol. 32, pp. 287–298, 2006.

[50] E. S. Siah, M. J. Sasena, J. L. Volakis, P. Y. Papalambros, and R. W. Wiese, "Fast parameter optimization of large-scale electromagnetic objects using DIRECT with Kriging metamodeling," *IEEE Transactions on Microwave Theory and Techniques*, vol. 52, pp. 276–285, January 2004.

[51] H. Stark and J. W. Woods, *Probability and Random Processes with Applications to Signal Processing*. New Jersey: Prentice Hall, 3 ed., 2002.

[52] M. J. Sasena, P. Y. Papalambros, and P. Goovaerts, "The use of surrogate modeling algorithms to exploit disparities in function computation time within simulation-based optimization," in *In The Fourth World Congress of Structural and Multidisciplinary Optimization*, 2001.

[53] D. D. Cox and S. John, "SDO : A statistical method for global optimization," in *in Multidisciplinary Design Optimization: State-of-the-Art*, pp. 315–329, 1997.

[54] A. G. Watson and R. J. Barnes, "Infill sampling criteria to locate extremes," *Mathematical Geology*, vol. 27, pp. 589–608, 14 March 1995.

[55] J. K. Yamamoto, "An alternative measure of the reliability of ordinary Kriging estimates," *Mathematical Geology*, vol. 32, pp. 489–509, May 2000.

[56] W. Ponweiser, T. Wagner, and M. Vincze, "Clustered multiple generalized expected improvement : A novel infill sampling criterion for surrogate models," in *Evolutionary Computation, 2008. CEC 2008. IEEE Congress on*, pp. 3515 – 3522, IEEE World Congress on Computational Intelligence, 1-6 June 2008.

[57] B. D. Youn, K. K. Choi, R.-J. Yang, and L. Gu, "Reliability-based design optimization for crashworthiness of vehicle side impact," *Structure and Multidisciplinary Optimization*, vol. 26, pp. 272–283, 2004.

[58] C. Audet, J. J. E. Dennis, D. W. Moore, A. Booker, and P. D. Frank, "A surrogate-model-based method for constrained optimization," in *the 8th AIAA/NASA/USAF/ISSMO Symposium on Multidisciplinary Analysis and Optimization*, 2000. Paper No. AIAA-2000-4891.

[59] K.-Y. Chan, S. J. Skerlos, and P. Y. Papalambros, "An adaptive sequential linear programming algorithm for optimal design problems with probabilistic constraints," *Journal of Mechanical Design*, vol. 129, no. 2, pp. 140–149, 2007.

[60] T. Zou, Z. P. Mourelatos, S. Mahadevan, and J. Tu, "An indicator response surface method for simulation-based reliability analysis," *Journal of Mechanical Design*, vol. 130, no. 7, pp. 070410–1, 2008.

[61] H.-H. Lee, *Taguchi Mrthod: Principles and Parctices of Quality Design.* Taipei, Taiwan: Gau Lih Book Co. Ltd., 2008.

自　　傳

國立政治大學圖書館

附錄一：中英對照表

Active Constraint 有效拘束條件

Best Linear Unbiased Predictor 最佳無偏估線性預測法

Correlation Length 相關距離

Cross Validation 交叉驗證法

Deterministic 定率

Expect Improvement 改善期望值指標

Expensive Function 黑盒子函數

Experimental Variogram 實驗變異圖

Feasible Region 可行解空間

Feasibility 可行解機率

Fitting 擬合

Geostatistics 地理統計學

Genetic Algorithm 基因演算法

Ill-Condition 不良條件

Infill Sampling Criterion 取樣指標

Initial Sample 初始樣本

Intrinsic Hypothesis 本質假設

Leave One Out 單點移出法

Limit State 極限狀態

Maximal Likelihood Estimation 最大似然方法

Maximum Variance 最大變異數指標

Nugget Effect 碎塊效應

Random Function 隨機函數

Influence Range 影響範圍

Semivariogram 半變異圖

Sill 尾檻值

Simulated Annealing Algorithm 模擬退火演算法

Spatial Correlation Function 空間相關函數

Stationary Hypothesis 平穩假設

Stochastic 序率

Theoretical Variogram 理論變異圖

Tunneling 隧道指標

Variogram 變異圖

Variogram Cloud 半變異雲圖