

Particle Swarm Guided Evolution Strategy for Real-Parameter Optimization

Chang-Tai Hsieh

NCLab Report No. NCL-TR-2006007

July 2006

Natural Computing Laboratory (NCLab)
Department of Computer Science
National Chiao Tung University
329 Engineering Building C
1001 Ta Hsueh Road
HsinChu City 300, TAIWAN
<http://nclab.tw/>

國立交通大學

資訊科學與工程研究所

碩 士 論 文

發展粒子引導式演化策略演算法以處理
實數參數之全域最佳化問題



Particle Swarm Guided Evolution Strategy
for Real-Parameter Optimization

研 究 生：謝長泰

指導教授：陳穎平 教授

中 華 民 國 九 十 五 年 六 月

發展粒子引導式演化策略演算法以處理實數參數之全域最佳化問題

**Particle Swarm Guided Evolution Strategy
for Real-Parameter Optimization**

研 究 生：謝長泰

Student：Chang-Tai Hsieh

指導教授：陳穎平

Advisor：Ying-Ping Chen

國 立 交 通 大 學

資 訊 科 學 與 工 程 研 究 所

碩 士 論 文



A Thesis

Submitted to Institute of Computer Science and Engineering

College of Computer Science

National Chiao Tung University

in partial Fulfillment of the Requirements

for the Degree of

Master

in

Computer Science

June 2006

Hsinchu, Taiwan, Republic of China

中華民國九十五年六月

發展粒子引導式演化策略演算法以處理實數參數之 全域最佳化問題

學生：謝長泰

指導教授：陳穎平

國立交通大學資訊科學與工程研究所

摘 要

遺傳演算法為一種模仿自然界演化現象所發展出的一套隨機搜尋的方法，在這些方法中，又以演化策略與粒子最佳化演算法為兩個最熱門的研究主題。這兩種方法皆為處理實數參數最佳化的問題，但其搜尋的行為卻迥然不同。藉由觀察演化策略與粒子演算法的搜尋行為，我們發現兩者皆有其優點與缺點。在演化策略中，起因於隨機擾動，突變運算子缺乏一個明顯的機制來引導搜尋行為至預期會有較佳解的區域，除此之外，個體在搜尋空間中的移動行為缺乏群體相互協調的機制，然而，其強大的篩選機制可將較好的個體保留至下一代；在粒子演算法中，其搜尋機制使用群體合作的概念，每一個個體在搜尋空間的移動行為將直接朝向預期會有較佳解的方向前進。本研究的目的試圖將演化策略與粒子演算法做一個概念層級上的結合，我們引進粒子演算法中，群體合作的機制，以此為概念來對於演化策略中偏向於隨機搜尋的突變運算子進行突變方向上的引導。本文將提出一個新的突變機制，稱之為引導式突變，結合引導式突變於傳統演化策略的搜尋架構下，進而發展出粒子引導式演化策略的最佳化方法。效能評估將建立在一組精心設計的測試平台下，由實驗結果得知所提的方法可達到良好的搜尋效能。

關鍵字：演化計算、演化策略、引導式突變、突變運算、粒子演算法、
群體智能

Particle Swarm Guided Evolution Strategy for Real-Parameter Optimization

student : Chang-Tai Hsieh

Advisor : Ying-Ping Chan

Institute of Computer Science and Engineering
National Chiao Tung University

ABSTRACT

Evolutionary algorithms are stochastic search methods that mimic the metaphor of natural biological evolution. Among of these algorithms, evolution strategy (ES) and particle swarm optimization (PSO) are two of the most popular research topics. Both of ES and PSO are deal with real-parameter optimization problems but have different search behaviors. By observing the search behavior of ES and PSO, we find both of them have strength and weakness. In ES, the mutation operator lacks an explicit mechanism to guide the search into promising direction due to random variance. Moreover, there is no coordination in the movement of individuals within the search space. However, the powerful selection procedure allows solutions with superior characteristics to pass these from generation to generation. In the PSO, the search mechanism used the swarm cooperation concept. Each particle will move toward the direction which is expected to be good. The objective of this article is tried to combine ES and PSO at the concept level. We introduce the concept of swarm cooperation of PSO into the mutation operator of ES for reducing the disturbance of mutation in the mutation direction. We proposed a new mutation operator called guided mutation. Combining the guided mutation into the traditional ES framework and develop a new optimization algorithm called particle swarm guided evolution strategy. Numerical experiments are conducted on a set of carefully designed benchmark functions and demonstrate good performance achieved by the proposed methodology.

Keyword : Evolutionary Computation , Evolution Strategy, Guided Mutation, Mutation Operator, Particle Swarm Optimization, Swarm Intelligence

誌 謝

本篇論文之得以順利完成，首先要向我的指導教授-陳穎平老師，致上最高的敬意，感謝老師在我的碩士生涯中耐心且專業的指導，讓我不僅在學業上得到正確的研究技能及知識，且學習到待人謙卑及處事積極的態度。

由衷感激孫春在老師、鐘雲恭老師、洪炯宗老師在論文口試期間不吝惠賜寶貴的意見，讓本論文之研究成果更趨於完善與流暢，而他們博學謙恭的態度，亦是我未來值得仿效的榜樣。

感謝自然計算實驗室成員柏甫、明昌、志宇、秉竹、宏偉、盈吟、文瑜等研究夥伴，以及欣諭、哲逸、智超、KIKI、明冠、大為、永發、智洋、東昇、書楷、士弘等多位好友。因為有大夥的鼓勵與協助，使得求學的路程，充滿溫馨及歡笑。

最後，感謝我的家人無私的奉獻與支持，給我勇往直前的力量，使我可以全力衝刺學業。還有，感謝我的可愛女友-韋如，因為她的陪伴與關懷，讓我的心滿溢著暖暖的幸福，讓我的生活不再是枯燥的 0 與 1。

藉此論文付梓之際，再次向協助過我的朋友們，致上最真切的謝意。

謝長泰 謹誌

民國九十五年六月三十日

於風城交大

目 錄

中文摘要	I
英文摘要	II
誌 謝	III
目 錄	IV
圖 目 錄	VI
表 目 錄	VII
第一章 緒論	1
1.1 研究動機與背景	1
1.2 研究目標	3
1.3 論文架構	4
第二章 背景知識與文獻探討	5
2.1 演化式計算簡介	6
2.2 演化策略	11
2.3 粒子演算法	24
2.4 相關文獻探討	26
2.5 本章總結	27
第三章 粒子引導式演化策略	28
3.1 想法起源	29
3.2 粒子演算法之引導概念	34
3.3 引導式突變	35
3.4 視覺化評估與觀察	41
3.4.1 引導式突變與傳統突變方式之比較	42
3.4.2 引導式突變與關聯式突變分析	47
3.5 引導式演化策略	51
3.6 本章總結	56
第四章 實驗結果與效能分析	57
4.1 實驗衡量標準	57
4.2 測試函式	59
4.3 實驗評比對象及參數設定	63
4.4 實驗結果與效能分析	64
4.4.1 引導式演化策略與傳統演化策略比較	64

4.4.2	引導式演化策略與其他進階演算法比較	72
4.5	本章總結	74
第五章	總結與未來工作	75
5.1	總結	75
5.2	未來工作	77
5.3	主要總結	78
附錄一	CEC'05 25 個 TEST FUNCTIONS 之特性與定義	80
附錄二	CEC'05 25 個 TEST FUNCTIONS 整體最佳解的個體目標變數與其適應值	92
附錄三	PSGES 所達到的最佳解之個體目標變數與其適應值	93
參考文獻		94



圖目錄

圖 2.1	演化計算通用示意圖	8
圖 2.2	(1+1)-ES 流程圖	11
圖 2.3	ES 基本演化流程	13
圖 2.4	單一突變強度之非關聯式突變	16
圖 2.5	n 維突變強度的非關聯式突變。	18
圖 2.6	關聯式突變。	20
圖 2.7	傳統演化策略流程圖。	21
圖 2.8	粒子演算法移動示意圖	25
圖 3.1	演化路徑示意圖	32
圖 3.2	引導式突變流程圖	38
圖 3.3	引導式突變與粒子移動行為比較	39
圖 3.4	傳統突變方式與引導式突變比較	40
圖 3.5	起始環境下個體的分布情形	42
圖 3.6	n 維突變強度的非關聯式突變演化示意圖	44
圖 3.7	關聯式突變演化示意圖	45
圖 3.8	引導式突變演化示意圖	46
圖 3.9	關聯式突變 ($\beta=20^\circ$) 演化示意圖	49
圖 3.10	關聯式突變 ($\beta=360^\circ$) 演化示意圖	50
圖 3.11	PSGES 概念示意圖	54
圖 3.12	引導式演化策略流程圖	55

表 目 錄

表 2.1	演化策略、演化式規劃、基因演算法比較表	9
表 2.2	演化式計算之優缺點比較表	10
表 2.3	策略參數設定與其所對應的個體表示法整理	23
表 4.1	25 個 test functions 特性分類.....	60
表 4.2	最佳解、搜尋範圍與起始範圍定義表	61
表 4.3	25 個 test functions 之求解臨界值定義.....	62
表 4.4	實驗參數設定	64
表 4.5	Results of Function 1-5	66
表 4.6	Results of Function 6-10	67
表 4.7	Results of Function 11-15.....	68
表 4.8	Results of Function 16-20	69
表 4.9	Results of Function 21-25	70
表 4.10	ES、PLES 與 PSGES 平均效能表現.....	71
表 4.11	PSGES 與其他進階演算法比較	73



第一章 緒論

1.1 研究動機與背景

現實世界中眾多的問題最終仍歸結為最佳化問題，最佳化問題存在於各個領域，對於其求解的方式，過去已有不少經典的演算法被提，對於某些特定的問題上確實能得到較好的結果。但這些傳統的方法對於目標函式有很大的限制，如連續、可微等。因此傳統方法對於問題的依賴性較強，以一個二元數學函式為例，一般需要用到目標函式一階或二階偏微分才可推導出來，而實際問題所遇到的目標函式往往更加複雜，有的甚至難以用數學公式表達出，因此，光是利用數值方法很難求得最佳解。

近年來快速發展的演化計算，為一種借鑑於達爾文進化論[1]而模擬出的一套隨機演算法，舉凡觀察自然、生物、與生命現象為手段，得到設計可應用於搜尋 (Search)、設計 (Design)、與最佳化 (Optimization)等各類問題之演算方法，皆屬於演化計算的研究領域。演化計算著眼於從一群(種族)潛在解(個體)中尋找問題的最佳解，透過在當前的潛在解之間反覆地進行如篩選、重組、突變等遺傳操作，來對於最佳解所在的位置逐步逼近。演化計算在於應用問題方面，至今已成功的運用在許多不同領域的最佳化問題上，如藥物設計、生物資訊、演化硬體、工業設計、電路設計、資料探勘、財經分析等。

演化策略(Evolution Strategy, ES)為演化計算的其中一個重要分支，其歷史可回溯至 1960 年代，德國柏林大學 Schwefel 與 Rechenberg 兩位研究學者為了解決工程上所遇到的最佳化問題進而發展出的一套最佳化演算法。在演化策略中，個體以一個實數向量來表示，最佳化的過程主要以突變的機制來進行，因此突變機制在 ES 中扮演著極為重要的角色，一個設計良好的突變機制將對於搜尋效能有決定性的影響，在 ES 的研究範疇裡，大部分的研究皆著重於突變機制下做探討與並加以改良以增進其效能。

另一方面，近年來興起的粒子演算法 (Particle Swarm Optimization, PSO)，由 Kennedy 和 Eberhart 於 1995 所提出，與 ES 同為實數參數最佳化的方法，然其搜尋上的行為卻迥然不同，PSO 的概念源自群體行為理論，啟發自觀察鳥群或魚群移動時，能透過個體間特別的訊息傳遞方式，使整個群體朝預期會有較佳解的方向而去，為模仿此類生物行為反應來尋求完成群體最大利益的方法。

經由對 ES 與 PSO 兩者的搜尋行為上做研究與探討，本論文的主要研究動機，著眼於對傳統 ES 的突變行為做改良，緣起於觀察出 ES 中在突變機制上的行為與限制，也就是突變行為以亂數決定所造成的不穩定性。故在本論文中，試圖擷取 PSO 中群體間相互合作的群體智能，以此為「概念」來對於偏向隨機搜尋的 ES 進行搜尋方向上的引導，賦予 ES 中個體突變的行為參照群體智能所獲得的資訊，進而達到更有具系統化的突變能力。

1.2 研究目標

本論文的主要目的為擷取 PSO 中群體智能的搜尋概念，將其在搜尋行為上的優點與傳統 ES 做概念層次上的結合，而非只是計算層次的結合。我們將重點著重在傳統 ES 上的突變機制並加以修改，發展一個利用個體間交互參照的訊息與群體於搜尋空間中所分布的行為與結構，使得 ES 在演化的效能上能以更直接有效的方式達到最佳化的結果，並且設法降低在突變的行為上以亂數決定所造成的不穩定性。

在我們的研究目標中，首先將提出一個結合 PSO 的搜尋概念，發展出一個在突變方向上具有引導性質的突變機制，我們將此稱作為引導式突變，接著將引導式突變與 ES 中原有的搜尋架構做整合，進而提出一套新的最佳化演算法，稱之為粒子引導式演化策略 (Particle Swarm Guided Evolution Strategy, PSGES)，最後以一系列公正的測試函式來評量所提 PSGES 的效能，並以傳統 ES 與其他進階演算法做比較，分析與探討本文所提的方法之效能與表現。

1.3 論文架構

本論文總共分成六個章節，其結構如下：

- 第一章為前言，此部份包含研究動機與背景、研究目標以及論文架構、說明本研究的重要性以及本研究所要達成的目標。
- 第二章為背景知識與文獻探討，此部份將對於與本研究有直接相關的背景知識作說明，其中包括 ES 與 PSO，並敘述近年來的相關研究。
- 第三章為本研究所提的 PSGES 之方法論，包含所提方法的想法起源，與傳統方法的差異以及對於所提的方法之詳細論述。
- 第四章為實驗結果與效能分析，此部份包含所要作為衡量標準的測試函式、所提方法的實驗結果，並與其他方法的比較以及效能上分析與討論。
- 第五章為本論文之結論以及未來工作。

第二章 背景知識與文獻探討

演化計算以觀察自然、生物、與生命現象為手段，得到設計可應用於搜尋、設計、與最佳化等各類問題之演算方法或解題技術的原則或啟發。ES 為演化式計算的一個重要分支，主要著重在實數參數最佳化的方法，最佳化的過程主要利用突變的機制來逐步朝向最佳解前進，因此突變機制在 ES 中扮演相當重要的角色。本文所提的 PSGES，將對於傳統 ES 的突變機制加以修改，目的為截取 PSO 中，群體間相互合作的概念行為，來增進 ES 中突變機制的效能，以期在個體突變的過程中，以更有系統的方式達到突變的效果。

在本章中，首先將針對與本研究相關的背景知識作說明，分成以下四個部份來闡述。第一節為演化式計算的基本簡介；第二節中，我們將對於 ES 的主要運作流程加以說明，其中包括編碼形式、重組、篩選、突變以及自動調節機制(self-adaptation)；第三節說明為 PSO 的運算模式；第四節為與本研究相關的文獻探討。

2.1 演化式計算簡介

相對於古典的最佳化的技術，演化式計算為一門相當年輕的最佳化演算法，主要為根據生物演化觀念所建構成的一種計算模式，利用達爾文進化論中「適者生存，不適者淘汰 (survival of the fittest)」的理論[1]，做為在廣大空間的搜尋機制，藉此可求解各種困難的組合最佳化問題，找到問題的最佳解。因此演化式計算可視為一個集合的統稱，依此模式下所發展出的演算法通稱為演化式演算法 (Evolutionary Algorithm, EA)，大致上可分類成三大類，其中包括基因演算法 (Genetic Algorithm, GA)[2]、演化式規劃 (Evolutionary Programming, EP)[3] 與演化策略 (Evolution Strategies, ES)[4]，三者之間的差異如表 2.1 所示[5, 6]。

族群式 (Population Based)的演化觀念為演化式計算最主要的特色，一個族群包含著許多個體，對於最佳化的問題上，個體經由特殊的編碼形式，單一個體就代表著對此問題上的一個潛在解(genotype)，個體則由基因變數的表現型(phenotype)來表現對於最佳化問題上的適應能力。擁有較好適應能力的個體將有更高的機會被保留並再次進行演化。仿照達爾文「物競天擇，適者生存」的精神，演化式計算則運用此想法來達到最佳化的目的，對於一個最佳化的問題，經由不斷的重組、突變、競爭與篩選，個體將逐漸地朝向問題的最佳解前進，圖 2.1 為演化式計算的通用流程圖。其中最主要的

運作機制有下列四點：

- 一、 重組： 在演化式計算中，重組機制 (Recombination)模仿自然界個體交配的過程，利用重組所產生的子代，將保留部分父代個體中的資訊。
- 二、 突變： 突變 (Mutation)為一個隨機擾動的行為，在最佳化的過程中，藉由將個體進行突變運算，來對個體的基因變數產生隨機的擾動，為造成個體彼此間差異的一個主要來源。
- 三、 競爭： 源自於達爾文「物競天擇」的觀念，個體間彼此相互競爭，經由計算 (Evaluation)個體的適應能力來評判個體好與壞的一個過程，擁有較佳適應能力的個體將有較高的機會被保留下來。
- 四、 篩選： 源自於達爾文「適者生存」的觀念，擁有較佳適應能力的個體將在篩選的機制 (Selection)下保留下來並成為下一代的人口。



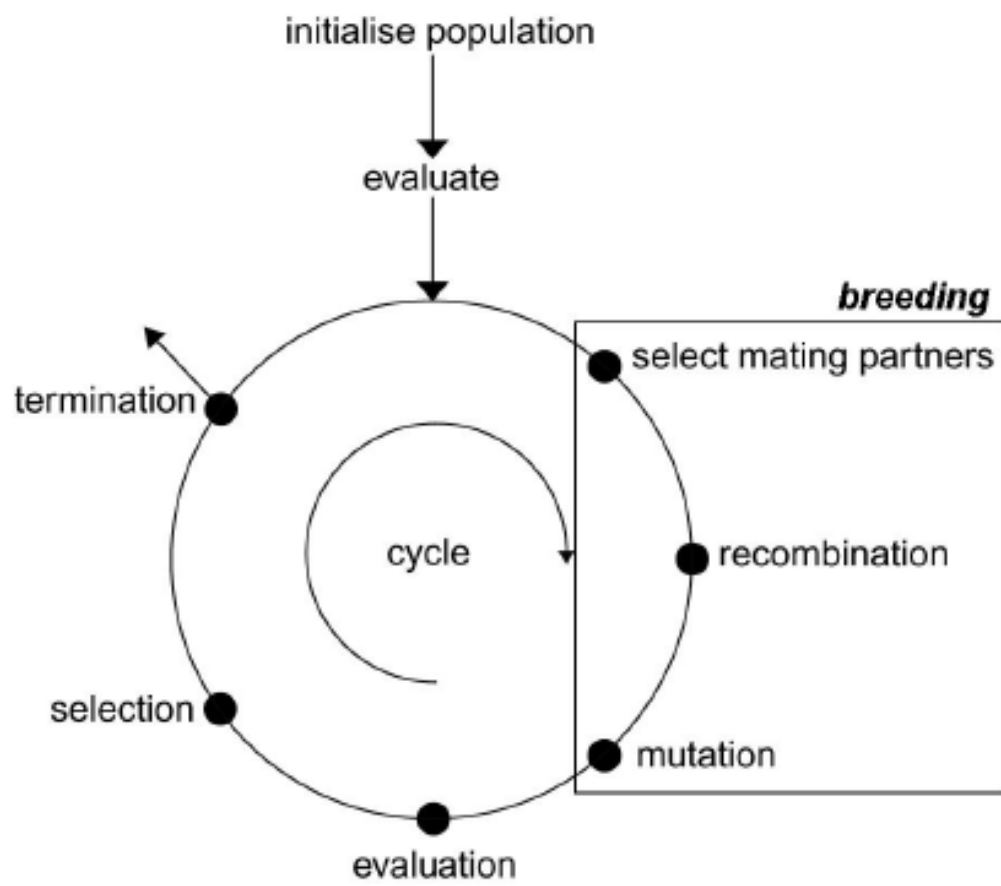


圖 2.1 演化計算通用示意圖

表 2.1 演化策略、演化式規劃、基因演算法比較表

	ES	EP	GA
Representation	Real-valued	Real-valued	Binary-valued
Fitness is	Objective function value	Scaled objective function value	Scaled objective function value
Self-Adaptation	Standard deviations and rotation angles	None (standard-EP), variances (meta-EP), correlation coefficients	None
Mutation	Gaussian, main operator	Gaussian, only operator	Bit-inversion, background operator
Recombination	Discrete and intermediate	None	z-point crossover, uniform crossover, only sexual, main operator
Selection	Deterministic, extinctive or based on preservation	Probabilistic, extinctive	Probabilistic, based on preservation
Constraints	Arbitrary inequality constraints	None	Simple bounds by encoding mechanism
Theory	Convergence rate for special cases, (1+1)-ES, (1+ λ)-ES, (1, λ)-ES, global convergence for (μ + λ)-ES	Convergence rate for special cases, (1+1)-EP, global convergence for (1+1)-EP	Schema processing theory, global convergence for elitist version

演化式計算的優點在於對各種類型的最佳化問題都可以表現出一定的求解能力，與傳統數學歸化法比較起來，演化式計算提供了較大的穩定性 (Stability) 與強健性 (Robustness)，尤其對於現實世界中較複雜的最佳化問題，例如多變數問題、高度非線性問題以及不可微分的問題來說，演化式計算都有較好的表現。然而，演化式計算的缺點之一，是只能找到一個近似的最佳解，而較難逼近到真正的最佳解上面。表 2.2 為演化式計算的優缺點比較。

在本論文提出的方法中，主要以演化式計算的一個重要分支：演化策略 (Evolution Strategies) 作為研究的主軸，並試圖將其與近年來興起的粒子演算法在搜尋的概念上做結合，因此在以下兩節中，我們將分別針對演化策略與粒子演算法的基本知識與相關背景作說明。

表 2.2 演化式計算之優缺點比較表

Advantage	Disadvantage
• No presumptions w.r.t. problem space • Widely applicable • Low development & application costs • Easy to incorporate other methods • Solutions are interpretable • Can be run interactively, accommodate user proposed solutions • Provides many alternative solutions • Suitable for changing conditions	• No guarantee for optimal solution within finite time • Weak theoretical basis • May need parameter tuning • Often computationally expensive

2.2 演化策略

1960 年代早期，德國人 Rechenberg 在柏林工業大學 (Technical University of Berlin, TUB)，為了解決流體力學中模型控制裡實數參數最佳化的問題，發展了第一個實驗性的演化策略方法，此為演化策略的開端。因為體認到工程上的一些非線性模型的數值問題，無法用傳統的數學線性方法求解。後來他與 Schwefel 合作發明了一種新的方法，即為「演化策略」，有關演化策略解決工程上非線性模型的數值問題其細節部份可參閱[7]。最早的演化策略非常簡單，類似隨機區域搜尋，稱為 (1+1)-ES，其演化流程如圖2.2所示。

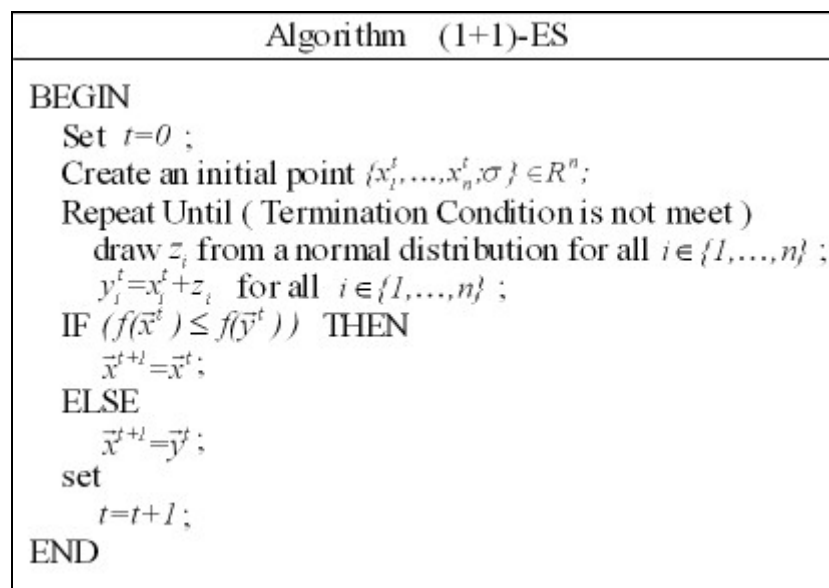


圖 2.2 (1+1)-ES 流程圖

ES 與其他遺傳演算法不同的地方，在於 ES 中個體的編碼是以實數 (real-valued) 的方式來表示，因此更適合處理工程上實數最佳化的問題。對於一個有 n 個參數需要被最佳化的問題，個體 $\vec{I}$ 直接以一個 n 維的實數向量 $\vec{x}$ 來表示，稱之為目標變數 (objective variables)。除了目標變數之外，每一個 $\vec{I}$ 的表示法中，還額外包含兩組策略參數 $\vec{\sigma}$ 與 $\vec{\alpha}$ ，其功能為調整目標變數中突變行為的依據，其中 $\vec{\sigma}$ 稱之為 step-sizes，用來調整個體的突變強度， $\vec{\alpha}$ 稱之為 rotation angles，用來調整個體的旋轉角度(或突變方向)。結合目標變數與策略參數，單一個體 $\vec{I}$ 的表示法如下所示[5]

$$\vec{I} = (\vec{x}; \vec{\sigma}; \vec{\alpha}) \in \mathbb{R}^n \times \mathbb{R}_+^{n_\sigma} \times [-\pi, \pi]^{n_\alpha}$$

其中 $n_\sigma \in \{1, \dots, n\}$, 代表 step-sizes 的個數； $n_\alpha \in \{0, (2n - n_\sigma)(n_\sigma - 1)/2\}$ ，代表 rotation angles 的個數。依據不同突變機制的採用，其個體的表示法也有所不同。ES 的基本演化流程與其他遺傳演算法類似，如圖 2.3 所示[5]。

```

t:=0;
initialize  $P(0):=\{\vec{a}_1(0), \dots, \vec{a}_\mu(0)\} \in I^\mu$ 
    where  $I = \mathbb{R}^n \times \mathbb{R}_+^{n_\sigma} \times [-\pi, \pi]^{n_\alpha}$ 
    and  $\vec{a}_k = (x_p, \sigma_j, \alpha_{j'}), i \in \{1, \dots, n\}$ 
         $j \in \{1, \dots, n_\sigma\}, j' \in \{1, \dots, n_\alpha\};$ 
    or  $\vec{a}_k = (x_p, \sigma_i), i \in \{1, \dots, n\}, j \in \{1, \dots, n_\sigma\};$ 
evaluate  $P(0): \{\phi(\vec{a}_1(0)), \dots, \phi(\vec{a}_\mu(0))\}$ 
    where  $\phi(\vec{a}_k(0)) = f(\vec{x}_k(0));$ 
while (termination criterion  $\iota(P(t)) \neq \text{ture}$ ) do
    recombine:  $\vec{a}'_k(t) := r'(P(t)) \quad \forall k \in \{1, \dots, \lambda\};$ 
    mutate:  $\vec{a}''_k(t) := m'_{\{\tau, \tau', \beta\}}(\vec{a}'_k(t)) \quad \forall k \in \{1, \dots, \lambda\};$ 
    evaluate:  $P''(t) := \{\vec{a}''_1(t), \dots, \vec{a}''_\lambda(t)\};$ 
         $\{\phi(\vec{a}''_1(t)), \dots, \phi(\vec{a}''_\lambda(t))\}$ 
        where  $\phi(\vec{a}''_k(t)) = f(\vec{x}''_k(t));$ 
    select:  $P(t+1) := \text{if } (\mu, \lambda)\text{-selection}$ 
        then  $s_{(\mu, \lambda)}(P''(t));$ 
        else  $s_{(\mu+\lambda)}(P(t) \cup P''(t));$ 
    t:=t+1;
end

```

圖 2.3 ES 基本演化流程

由圖 2.3 的 ES 基本演化流程可發現，執行最佳化的運作機制主要發生於重組 (recombination)、突變 (mutation)、篩選 (selection) 三個遺傳操作，其中又以突變的遺傳操作最為重要，以下針對這三種遺傳操作加以說明。

重組 (Recombination)

重組的遺傳操作發生在至少兩個或兩個以上的個體，換句話說，在族群式的演化策略(population-based evolution strategy) $\mu > 1$ 的情況下才可進行重組，一次重組產生一個新的個體。重組的主要目的，為交換個體彼此間資訊，以達到資訊流通的效果。一般的作法為先設定一個大於 1 的 ρ 值作為參與重組的父代個數，稱之為候選者(candidates)。演化策略有許多變形的重組方式，Bäck[5]列出了下列五種較常使用的重組運算，如下所示

$$x'_i = \begin{cases} x_{a,i} & \text{no recombination} \\ x_{a,i} \text{ or } x_{b,i} & \text{discrete} \\ (x_{a,i} + x_{b,i})/2 & \text{intermediate} \\ x_{a_i,i} \text{ or } x_{b_i,i} & \text{global discrete} \\ (x_{a_i,i} + x_{b_i,i})/2 & \text{global intermediate} \end{cases} \quad (1.1)$$

其中 x'_i 為經重組後的子代第 i 個變數的值， $x_{a,i}$ 與 $x_{b,i}$ 代表從 ρ 個父代中隨機挑選出兩個父代個體的第 i 個變數的值。

- no recombination : x'_i 的值由 $x_{a,i}$ 取得，無重組的動作發生。
- discrete : x'_i 的值隨機由 $x_{a,i}$ 或 $x_{b,i}$ 來產生。
- intermediate : x'_i 的值由 $x_{a,i}$ 與 $x_{b,i}$ 的平均值來決定。
- global discrete : 與 discrete 類似，但每個變數都需重新挑選兩個父代。
- global intermediate : 與 intermediate 類似，但每個變數都需重新挑選兩個父代。

重組運算可同時作用於目標變數與策略參數，且可分別使用不同的重組方式。一般來說，目標變數使用 discrete 或 global discrete 的重組方式；策略參數 σ 使用 intermediate 或 global intermediate 的重組方式；策略參數 $\bar{\alpha}$ 的部份則偏向以 no recombination 的方式來進行。

突變 (Mutation)

突變為 ES 中最重要的運算，也是達到個體間差異的主要來源。相對於重組運算利用已存在個體之間的資訊，彼此交換進行訊息的流通，突變則是利用隨機擾動的方式產生一個含有新的資訊之個體。突變運算依據策略參數個數的不同，將直接影響到個體所產生的突變行為。一般來說，對於一個 n 維的問題上，我們可將個體的突變分成三個類型來討論：

單一突變強度之非關聯式突變 (Uncorrelated Mutation with One Step-Size)

在此突變機制下，策略參數的設定為 $n_\sigma=1$ ， $n_\alpha=0$ ，也就是說利用單一 step-size 來調整目標變數裡的每一個變數，因此個體的表示法如下

$$\vec{I} = (x_1, \dots, x_n; \sigma)$$

個體突變方式分成兩個步驟，第一先對策略參數 σ 進行突變，接著再利用突變後的 σ 來調整目標變數 $\vec{x}$ 裡的每一個變數，其公式如下

$$\begin{aligned}\sigma^{(t+1)} &= \sigma \cdot e^{\tau \cdot N(0,1)} \\ x_i^{t+1} &= x_i^t + \sigma^{(t+1)} \cdot N_i(0,1)\end{aligned}\tag{1.2}$$

其中 $N_i(0, 1)$ $i \in 1, \dots, n$ 是平均值為 0，標準差為 1 的常態分佈隨機變數 (normal distributed random variable)。 $\tau \propto 1/\sqrt{n}$ 代表著 σ 的學習比率 (Learning Rate)。在這種突變機制下，我們只使用一個策略參數 σ 來對目標變數的每一個變數來做調整，因此突變後的子代在解空間上的分布範圍將形成一個以父代為中心的圓形區域。如下圖 2.4 所示

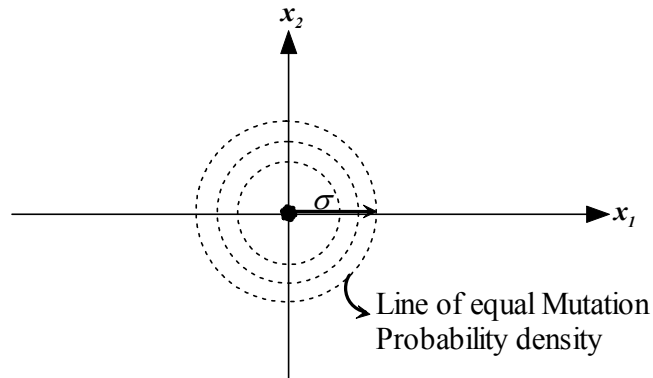


圖 2.4 單一突變強度之非關聯式突變

n 維突變強度之非關聯式突變 (Uncorrelated Mutation with n Step-Sizes)

相對於利用單一 step-size 來調整目標變數的每一個變數 $x_i, i \in 1, \dots, n$ ，在某些情況下，我們希望對於目標變數的每一個變數 $x_i, i \in 1, \dots, n$ 都單獨地使用獨自的 step-size 來調整，因此策略參數的數目設定增加成 $n_\sigma = n$ ， $n_\alpha = 0$ ，個體的表示方式如下

$$\vec{I} = (x_1, \dots, x_n ; \sigma_1, \dots, \sigma_{n_\sigma})$$

整體的突變機制有些許的變動，我們需要對 $\vec{\sigma}$ 中每個維度 $\sigma_i, i \in 1, \dots, n$ 進行突變，在利用突變後的 σ_i 來調整目標變數的第 i 個變數值，如下所示

$$\begin{aligned} \sigma_i^{t+1} &= \sigma_i^t \cdot e^{\tau' \cdot N(0,1) + \tau \cdot N_i(0,1)} \\ x_i^{t+1} &= x_i^t + \sigma_i^{t+1} \cdot N_i(0,1) \end{aligned} \quad (1.3)$$

其中 $N_i(0, 1) i \in 1, \dots, n$ 是平均值為 0，標準差為 1 的高斯分佈隨機變數， $\tau' \propto 1/\sqrt{2n}$ 與 $\tau \propto 1/\sqrt{2\sqrt{n}}$ 分別代表全域學習率 (Global Learning Rate) 與區域學習率 (Local Learning Rate)，在 $n_\sigma = n$ 的情況下，由於使用個別的 step-size $\sigma_1, \dots, \sigma_n$ 來調整其對應的目標變數 $x_1, \dots, x_n$ ，因此在搜尋空間上，突變後的子代將形成以父代為中心的橢圓形區域。如圖 2.5 所示。

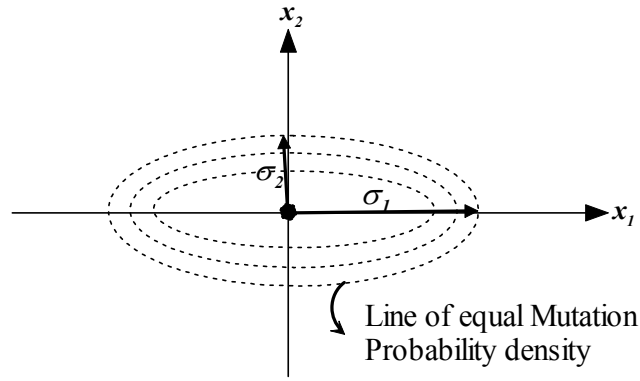


圖 2.5 n 維突變強度的非關聯式突變。

關聯式突變 (Correlated Mutation)

比較 $n_\sigma=1$ ， $n_\alpha=0$ 與 $n_\sigma=n$ ， $n_\alpha=0$ 兩種情況，雖說利用 n 個 step-sizes $\sigma_1, \dots, \sigma_n$ 來調整各自調整目標變數 $x_1, \dots, x_n$ ，遠比只利用單一 step-size 來調整目標變數較為有效益。但在 $n_\sigma=n$ 的情況下仍然存在著些許的缺點，也就是個體可突變的範圍，受限於座標軸系統造成突變橢圓與座標軸系統相依的缺點，1981 年 Schwefel[8]提出了關連性突變的機制，其目的為希望突變的橢圓區域能夠達到與座標軸系統獨立效果，此突變機制需要加入第二個策略參數 α 。策略參數個數的設定變成 $n_\sigma=n$ ， $n_\alpha = n(n+1)/2$ ，個體的表示法如下

$$\vec{I} = (x_1, \dots, x_n; \sigma_1, \dots, \sigma_{n_\sigma}; \alpha_1, \dots, \alpha_{n_\alpha})$$

突變的機制將變成先突變策略參數 σ ，再突變策略參數 α ，最後再對目標變數 $\vec{x}$ 進行突變，其數學表示式如下

$$\begin{aligned}
\sigma_i^{t+1} &= \sigma_i^t \cdot e^{\tau' \cdot N_i(0,1) + \tau \cdot N_i(0,1)} \\
\alpha_j^{t+1} &= \alpha_j^t + \beta \cdot N_j(0,1) \\
\vec{x}^{t+1} &= \vec{x}^t + \vec{N}(0, C(\vec{\sigma}, \vec{\alpha}))
\end{aligned} \tag{1.4}$$

其中 $N_i(0, 1) \quad i \in 1, \dots, n$, $N_j(0, 1) \quad j \in 1, \dots, n(n-1)/2$ 是平均值為 0 , 標準差為 1 的高斯分佈隨機變數 ; $\tau' \propto 1/\sqrt{2n}$ 與 $\tau \propto 1/\sqrt{2\sqrt{n}}$ 分別代表著 $\vec{\sigma}$ 的全域學習率 (Global Learning Rate) 與區域學習率 (Local Learning Rate) , $\beta \approx 0.0873 (5^\circ)$ 為控制 $\vec{\alpha}$ 的學習率 ; $\vec{N}(0, C(\vec{\sigma}, \vec{\alpha}))$ 為平均值為 $\vec{0}$, 標準差為協方差矩陣 (covariance matrix) C 的常態分佈所取樣的隨機變數向量 , 協方差矩陣 C 由突變後的 $\vec{\sigma}$ 與 $\vec{\alpha}$ 所組成 , 其元素如下



$$\begin{aligned}
c_{pq} &= \sigma_i^2, \\
c_{pq, p \neq q} &= \begin{cases} 0 & \text{no correlations,} \\ \frac{1}{2}(\sigma_p^2 - \sigma_q^2) \tan(2\alpha_{pq}) & \text{correlations.} \end{cases}
\end{aligned} \tag{1.5}$$

其中 , $c_{pq} \quad (p \in \{1, \dots, n-1\} , \quad q \in \{p+1, \dots, n\})$ 藉由旋轉角度向量 α_k ($k = \frac{1}{2}(2n-p)(p+1) - 2n+q$ [9]) 來表示可將一個非關聯性的突變向量 (uncorrelated mutation vector) 轉換成一個具有關聯性的突變向量 (correlated mutation vector) 。且能使得協方差矩陣具有正定的性質 , 並使得突變的橢圓在座標軸上朝任意方向旋轉[9] , 圖 2.6 為個體突變行為的示意圖。

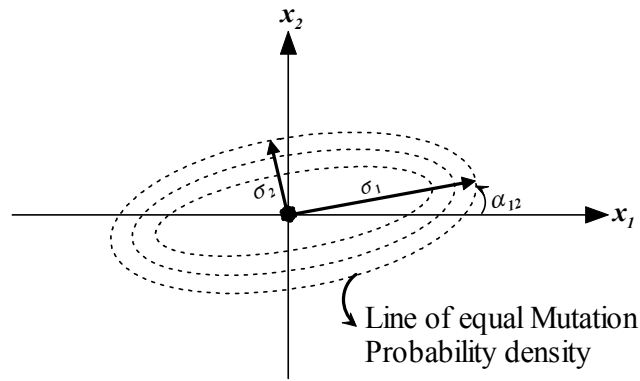


圖 2.6 關聯式突變。

篩選(Selection)



在 ES 中，篩選的運算可分為兩個主要的部份，一個為父代篩選(Parent Selection)，另一個為倖存者篩選 (Survivor Selection)。所謂父代篩選，起因於重組運算中，我們需要從 ρ 個候選者當中隨機挑選兩個至多個個體並參與重組運算，在這些候選者當中，每個個體將有相同的機率被選上。倖存者篩選又可分成兩個部份， (μ, λ) -ES 及 $(\mu+\lambda)$ -ES，兩者主要差別在於 $(\mu+\lambda)$ -ES 會將父代及子代族群一併進行評估；而 (μ, λ) -ES 會忽略父代資訊，而只考慮子代族群進行評估，再將進行評估的集合中所有個體依適應值排序，並挑選較佳的 μ 個體成為新的父代族群，進入下一代演化程序。

圖 2.7 為傳統演化策略的示意圖。

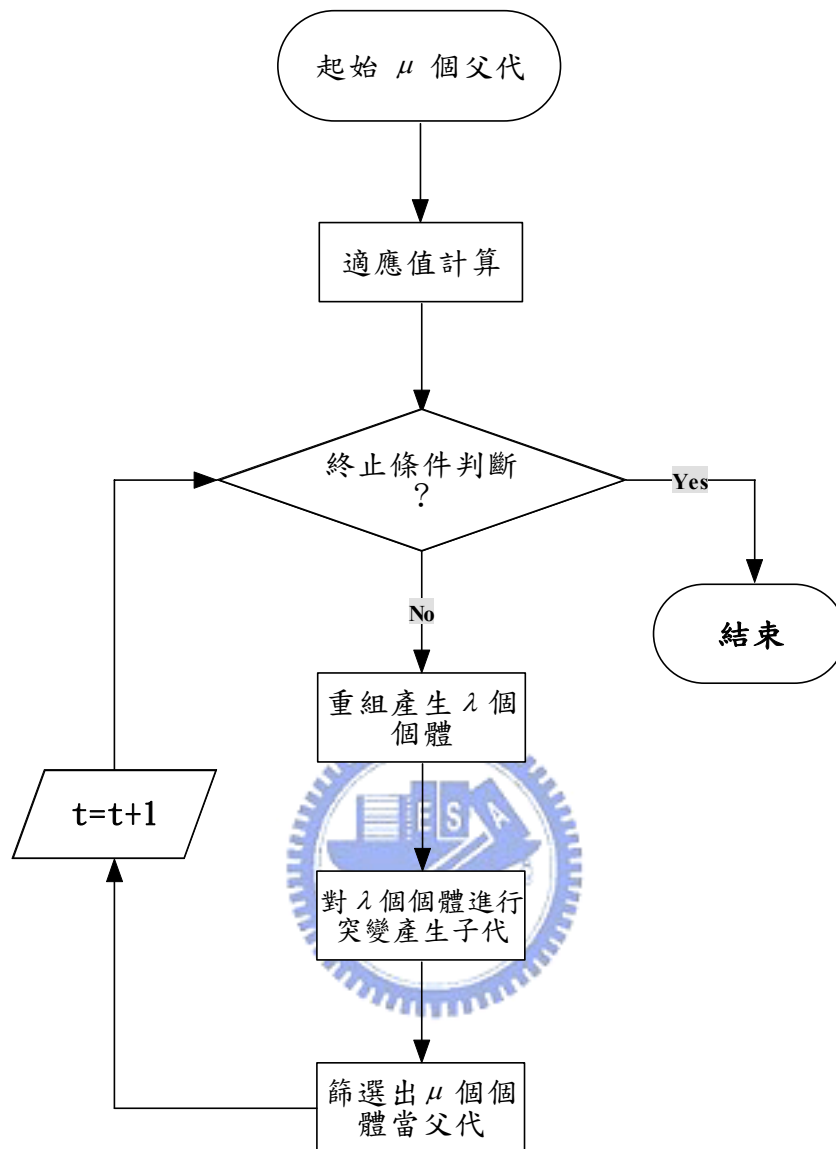


圖 2.7 傳統演化策略流程圖。

自動調節機制 (self-adaptation)

在討論完 ES 的基本運作模式後，接者我們將重點放在 ES 最具特色的一個機制：自動調節機制(self-adaptation)。

Self-adaptation 為演化策略中主要的一個特性，此機制賦予個體在最佳化的過程當中，自動調節個體的突變行為至最佳的狀態，作法為將策略參數依附在目標變數共同演化的方式來達成。在 self-adaptation 的想法中，經由先突變策略參數及再突變目標變數的方式，若先對策略參數進行突變後產生所謂好的策略值，將會直接地反映在目標變數的突變行為上，因而產生一個較好的子代個體。策略參數的品質好壞，則是間接地經由突變後產生子代的適應值(fitness value)來衡量。

最佳化的過程中，若能迅速的調整策略參數達到所謂較好的設定值，對其搜尋效能將有顯著的提升，在突變的架構下，影響 self-adaptation 調整品質及速度的關鍵，主要為參與調整的策略參數個數 n_σ 與 n_α ，表 2.3 歸納了三種最常使用的策略參數設定與其所對應的個體表示法。在策略參數個數 $n_\sigma=1$ ， $n_\alpha=0$ ，也就是單一突變強度之非關聯式突變的情況下，由於參與調整的策略參數只有一個，因此，self-adaptation 能迅速且有效地將 step-size 調整至適合的強度；在 $n_\sigma=n$ ， $n_\alpha=0$ 的 n 維突變強度之非關聯式突變地情況下，參與調整的策略參數個數與問題維度呈線性的成長，self-adaptation 對

策略參數的調整將比單一突變強度之非關聯式突變慢了許多，不過由於其較具彈性的突變能力(每一個維度以獨自的 step-size 調整)，故其仍然可達到較好的搜尋結果；最後一個也是最複雜的一種突變方式， $n_\sigma=n$ ， $n_\alpha=n(n-1)/2$ 的關聯性突變情況下，參與調整的策略參數在 n_σ 的部份與問題的維度呈線性成長，在 n_α 的部份將與問題的維度呈現二次平方項次型(quadric)的成長，即使在一個較低維度的問題上，對於所要調整的 rotation angles 卻已大大地超越 step-sizes 與目標變數所要調整的總和個數，因而 self-adaptation 將導致緩慢的收斂現象，此現象也是關連性突變機制不被廣泛使用的主要原因，即使其突變的能力上較前兩者更具彈性，本論文的研究重點將特別針對此現象加以深入研究與探討，進而提出一個利用群體智能的引導機制來對其作改良。



n_σ	n_α	Structure of individuals	Remark
1	0	$(x_1, \dots, x_n; \sigma_{n_\sigma})$	Uncorrelated Mutation
n	0	$(x_1, \dots, x_n; \sigma_1, \dots, \sigma_{n_\sigma})$	Uncorrelated Mutation
n	$n(n-1)/2$	$(x_1, \dots, x_n; \sigma_1, \dots, \sigma_{n_\sigma}; \alpha_1, \dots, \alpha_{n_\alpha})$	Correlated Mutation

表 2.3 策略參數設定與其所對應的個體表示法整理

2.3 粒子演算法

由於本論文所提出的方法，主要為擷取粒子演算法(Particle Swarm Optimization, PSO)搜尋的概念[10, 11]來對於 ES 中的突變機制加以修改，在提出本文的方法之前，我們將對 PSO 的基本特性以及其搜尋的主要精神做描述，對於 PSO 的相關變形以及細節部份在[12]有完整的介紹。

PSO 的概念源自群體行為理論，啟發自觀察鳥群或魚群行動時，能透過個體間特別的訊息傳遞方式，使整個團體朝預期會有較適合的方向、目標而去，此概念為模仿生物行為的反應來尋求完成群體最大利益的方法。PSO 的基本演算模式如下：首先以均勻分佈(uniform distribution)隨機產生初始粒子群，每一個粒子都可視為所求解問題上的一個潛在解，粒子群會參考粒子本身的最佳經驗，以及整個群體的最佳經驗，選擇修正其移動方式，重複不斷地進行修正之後，預期粒子群將會逐漸朝向最佳解前進。PSO 的主要運算公式可用兩個的式子表示：

$$\vec{v}^{t+1} = \vec{v}^t + c_1 \cdot rand_1^t() \cdot (\vec{x}_{pbest}^t - \vec{x}^t) + c_2 \cdot rand_2^t() \cdot (\vec{x}_{gbest}^t - \vec{x}^t) \quad (1.6)$$

$$\vec{x}^{t+1} = \vec{x}^t + \vec{v}^{t+1} \quad (1.7)$$

其中 $\vec{v}$ 為粒子移動的速度， t 為演化的代數， c_1 與 c_2 為兩個大於零的控制參數， $rand()$ 為一個介於[0,1]的常態分布隨機變數， $\vec{x}$ 代表粒子目前在解空

間上的位置， $\bar{x}_{pbest}$ 為粒子本身所達到的最佳解， $\bar{x}_{gbest}$ 為整群粒子群所找到的最佳解。依據此公式的計算方式，每一個粒子將會向 $\bar{x}_{pbest}$ 及 $\bar{x}_{gbest}$ 合成向量的方向移動，以增加改善 $\bar{x}$ 的機會，同時又設計了隨機(stochastic)的干擾量，以避免掉入局部最佳解(local optimum)。因此，粒子的向量修正原則是取得鑽探搜尋 (exploitation search) 以及探索搜尋 (exploration search) 的平衡，符合現代經驗法則(modern heuristics)之精神。由於實作簡單又具良好的效能，至今已成為廣為熟知的最佳化技術。

圖 2.8 為粒子演算法中粒子移動行為的示意圖，每個粒子借由參照本身所達最佳解與群體所達最佳解的位置，決定下一步所前進的方向與移動的強度，我們若以整個群體的觀點來觀察粒子們移動的行為，會發現粒子群們在移動的過程中，逐漸地朝向目前所找到的最佳解 $\bar{p}_g$ 的方向再度聚集。

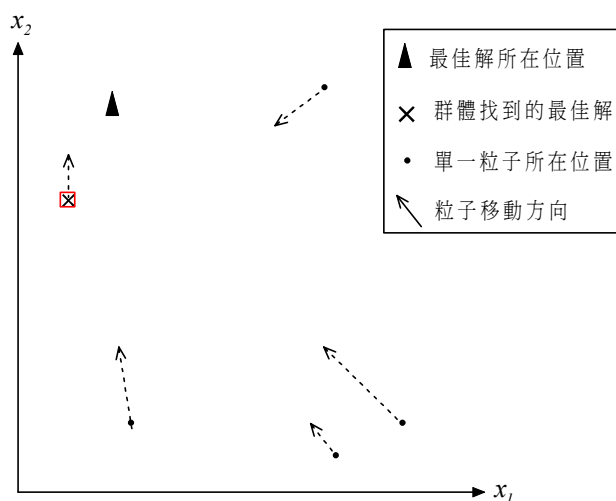


圖 2.8 粒子演算法移動示意圖

2.4 相關文獻探討

在這一節中，我們將對於近年來與本論文有關聯的相關研究做一個整理與描述。ES與其他演化演算法不同之處，在於其研究發展完全建立在嚴密的數值分析與機率理論上，對於其細節部份可參閱[13-19]。

起因於在突變運算上造成的過度隨機所產生的不穩定性，Ostermeier於1994年提出derandomized approach[20]的想法來增進self-adaptation的效能。Hansen於1996年引進了evolution path的概念[21]，利用過去所經歷過的路徑，來對於ES中個體下一步的突變機率分佈進行預測，並於2001年發展出一套completely derandomized approach的方法[22]，使用真正作用於目標變數的變異量來對策略參數進行調整。

Hildebrand於1999年提出了Directed Mutation機制[23]，其主要的想法為利用一個非對稱性的突變機率分佈，使所產生的子代有較高的機率往適合的方向進行突變。Zhang於2005年結合GA與EDA兩種不同的運算機制，利用EDA所維持的適當解分布之機率模型中，作為一個全域搜尋的資訊，進而提出了guided mutation的想法，來引導GA的突變行為。Miranda於2002年提出了EPSO的方法[24, 25]，在EPSO的想法裡，作者主要觀察到ES中的個體之間，缺乏合作的機制但卻有強大的突變運算(self-adaptation)以及篩選能力，PSO則是以群體合作的方式進行最佳化，但卻在對控制參數的調整上顯的較為薄弱，因而將ES與PSO做結合，其精神為利用ES領域的Self-Adaptation機制，來對PSO中的所使用到的控制參數做調整。

2.5 本章總結

在本章節中，我們首先介紹了傳統 ES 的歷史背景以及 ES 中相關的里程碑。接著說明 ES 的演化機制以及重要組成，其中包括編碼形式、重組、突變等重要的運算元。本論文中所提出的方法，與 PSO 有直接性的相關，因此，對於 PSO 的運作機制及其特性做說明。最後的相關研究中，描述了近年來 ES 的發展，並對與本論文有相關的研究做一概略性的描述。



第三章 粒子引導式演化策略

ES 至 1964 年被提出以來，許多改良的方法也陸續地被提出用以增進 ES 的演化效能，本章節將對於我們提出的方法，稱作粒子引導式演化策略 (Particle Swarm Guided Evolution Strategies, PSGES) 做一個完整性的介紹。PSGES 為針對實數問題上，所發展出的一套最佳化演算法，其主要精神為將傳統的 ES 與目前較新穎的 PSO 做一個本質上及概念上的結合。目的為擷取 PSO 中群體間資訊流通的機制，賦予 ES 中個體的突變運算達到更有效率的突變行為。

以下將分節來對我們提出的方法作一個完整性的介紹，我們分成六個小節討論：

- 第一節裡，我們將先敘述 PSGES 的想法起源，探討突變在搜尋行為上的限制，並解釋如何加以修改來增進其搜尋效能。
- 第二節中，說明如何運用 PSO 方向引導的概念與偏向隨機搜尋的突變行為在方向上做結合。
- 在第三節裡，我們將介紹一個新的突變機制，稱之為引導式突變(guided mutation)，在本節中我們將說明如何修改傳統 ES 中的突變機制，使其在突變的過程中具有引導的性質，以增進突變的效能。
- 第四節以一個視覺化的實驗模擬並驗證引導式突變的搜尋效益，並與傳

統的突變機制做比較。

- 第五節我們將引導式突變與傳統 ES 原有的搜尋架構做整合，完整的描述 PSGES 的運作流程，並討論可能影響其效能的相關因素與組成。
- 最後以第六節做為本章的總結。

3.1 想法起源

回溯至 1960 年代，Schwefel 與 Rechenberg 以實驗性的方式作為啟發，並提出 ES 以來，至今已有許多改良的方法陸續被提出，觀察 ES 的演化行為，我們發現突變機制為演化策略中造成個體間差異的最主要來源，因此一個設計良好的突變機制，將直接影響到 ES 的搜尋效能。近年來許多相關研究皆為針對突變的機制著手[20, 26, 27]，以期在突變的過程中更進一步地增進其搜尋的效能。在我們的研究中，經由觀察傳統突變在搜尋行為上的現象及其限制，因而提出一種新的突變機制，稱之為引導式突變(guided mutation)，在介紹引導式突變之前，首先我們將對於傳統演化策略的突變機制在搜尋行為上的現象作一番整理。

在傳統 ES 的架構下，策略參數為突變變異量調整的依據，根據不同的突變行為而採用不同的策略參數設定，經由突變產生多個子代，利用篩選的機制(Selection Operator)從突變後的子代中挑選一些較適合的解，擁有較

好適應值的個體將與以保留，並成為下一代的人口。在此架構下，促使群體往較好方向前進的動力，主要利用篩選機制來達成，在我們提出的引導式突變的想法中，希望直接在突變的運算上，藉由參考某些啟發性的資訊來引導個體往預期會有較佳解的方向產生突變，目的為避免將搜尋資源花費在無效的突變運算上，進而有效的減輕篩選機制在挑選適合個體上的負擔。

回顧第二章第二節，在關聯式突變的機制下，控制個體突變行為的策略參數分成兩部份：step-sizes σ 與 rotation angles α 。Step-sizes σ 用來調整個體向外擴展的突變強度；rotation angles α 則是用來調整個體在搜尋空間上突變橢圓的旋轉角度，以此來決定個體的突變方向，個體將可達到最具彈性的突變行為(突變橢圓與座標軸系統獨立的特性)。突變首先作用於策略參數 σ ，用以調整個體的突變強度，接著再對於策略參數 α 進行突變以決定個體的突變方向，然而由於參與調整策略參數 $n_\alpha = n(n-1)/2$ 與所要求解問題之維度呈現二次平方項次的成長，在決定個體突變方向上，若以 self-adaptation 機制運作，將不能夠立即且有效地將其做調整，故在突變方向的決定上仍然存在著諸多不穩定的現象。在目前相關的研究當中，已有許多理論證明 self-adaptation 在 $n_\sigma=1, n_\alpha=0$ 與 $n_\sigma=n, n_\alpha=0$ 兩種形式的突變機制下，能夠有效調整突變強度至適當的強度大小，並對於其收斂的行為做理論上的分析探討[28-30]；然而自動調節機制同時作用於 σ 與 α 兩個部份

(關連性突變)，起因分析及計算上的複雜度，至今仍缺乏有力的理論證明[31, 32]。歸納其原因如下：

1. 起因於突變的運算機制，為了達到有效自動調節機制同時作用於突變強度與旋轉角度上，父代的人口數(μ)也必須要大幅度的增加，對於一個 n 維的問題， $\mu \approx n^2$ 才能在合理的演化世代數中有效地調整旋轉角度至適當的值，在一般最常用的設定如 $\mu=15$ ， $\lambda=100$ 的情況下，明顯的不能達到此要求[9, 33]。
2. 起因於關聯性突變的運算機制，旋轉角度與突變強度有密切的關連，也就是說，對於所謂適合的旋轉角度設定將不能夠獨立地被判斷並做調整 [22]。
3. 篩選機制在挑選適合的策略參數上將會造成錯誤判斷的情況。在關連性突變下，突變強度與旋轉角度同為可被調整的策略參數，在某些情況下，若個體有較好的表現，我們不能明確的指出是突變強度有較好的設定亦或是旋轉角度有較好的設定，篩選機制在策略參數層級的篩選上將會有高度擾動的情況產生[22]。

基於以上所描述的原因，關聯性突變的突變行為與座標軸獨立的優點也相對的打折扣。我們將重點放在關聯性突變的旋轉角度上，在傳統的作法裡，對於關聯性突變的旋轉角度設定，主要還是以實驗性所得的經驗法則[34]，利用 5° 的角度量來做為高斯分佈的標準差，並對其做隨機的擾動以決定個體下一步的旋轉角度，此方式須花費大量的計算成本來對於角度做調整，搜尋的資源將花費在無效的突變行為上。是故，我們可以說關聯性突變在方向的決定上，仍缺乏一種明顯且有效的機制來促使個體往適合的方向進行突變。

起因於利用亂數來決定個體突變行為上所造成的不確定，因此有許多不同的方法被提出以改良突變在方向決定上所造成的不穩定現象，2001 年 Hansen 與 Ostermeier 提出了一種 derandomized 的方法[22]，其想法為利用個體過去演化所經歷過的演化路線(Evolution Path)，來預期最有可能發生最佳解的方向，以此為根據來判定個體下一步可突變的方向與強度，如圖 3.1 所示。

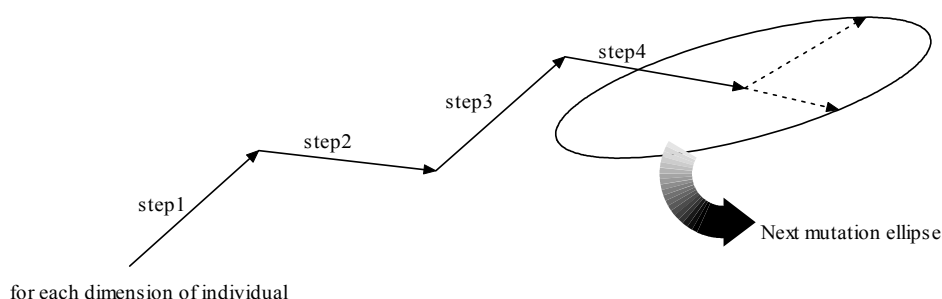


圖 3.1 演化路徑示意圖

在我們的想法中，同樣緣起於 self-adaptation 在大量旋轉角度的策略參數需做調整的情況下，造成個體突變方向上的不穩定現象，因此在旋轉角度 $\bar{\alpha}$ 的調整，試圖引入另一種啟發式 (heuristics) 的資訊以更具決定性的方式來對旋轉角度上進行設定，而不以 self-adaptation 的方式來做調整。此外，在 step-size 的部份，仍然保留使用 self-adaptation 來對其做調整。我們的作法為對傳統突變機制加以修改，進而設計出一個新的突變機制，稱之為引導式突變。引導式突變設法在旋轉角度的決定上消除其隨機擾動的性質，更確切的說，我們希望保留個體的突變橢圓與座標軸獨立的特性，並以更明確的方向引導資訊，希望在突變的機制下直接促使個體朝向預期會有較佳解的方向前進。

有鑑於此，我們從傳統 ES 中，利用關聯性突變原有的機制出發，試圖尋找一種不需人為介入且有效的輔助資訊來引導個體的突變行為，進而增進 ES 的突變效能。在下一節裡我們將說明如何利用 PSO 之群體智能的引導觀念，來做為方向引導上的輔助資訊，並說明如何與 ES 在「概念層次」上做結合。

3.2 粒子演算法之引導概念

如上一節所提到，在關聯性突變的機制下，對於旋轉角度上的決定，起因於亂數隨機地變動，因此個體在突變的行為上缺乏一個明顯的引導機制來促使群體往較佳解的方向前進，在這一節中，我們將說明如何擷取 PSO 原有的「搜尋概念」，作為引導式突變中方向引導的資訊，並解釋如何與 ES 在搜尋行為上做結合。

如第二章第三節所提到，PSO 為近年來興起的一套實數參數最佳化演算法，與 ES 比較，兩者皆為處理實數參數問題上所發展出來的最佳化方法，然而其在搜尋行為上卻有顯著的不同。在 PSO 的搜尋行為中，粒子在解空間上移動的方式相較於 ES 來說，在方向的決定上較具有明確的決定性。同樣地，我們也可將決定粒子移動的行為分成兩個部分討論，第一為移動強度的部分，第二為移動方向的部分。我們主要針對 PSO 在移動方向的部份討論，對於移動強度的探討可參閱[35, 36]。在 PSO 的粒子移動的過程裡，每一個粒子所前進的方向決定於自身所達的最佳解 $\vec{p}_i$ 與全體所達到的最佳解 $\vec{p}_g$ ，若以整個群體的觀點來觀察粒子們移動的行為，會發現粒子群們在移動的行為模式，會逐漸地朝向目前所找到的最佳解 $\vec{p}_g$ 的方向聚集。值得一提的是，對於一個最佳化的問題上，我們並不知道其真正最佳解所在的位置，因此在粒子演算法尋找最佳解的「概念」上，我們可視

$\bar{p}_g$ 的所在位置當作目前最有可能產生最佳解的區域。因此，粒子在移動的過程中，依據其計算模型，粒子的移動行為將逐漸地被 ”引導” 至 $\bar{p}_g$ 所在的方向。分佈於搜尋空間上的粒子，在此移動的過程中，若搜尋到另一個比 $\bar{p}_g$ 還要好的解 $\bar{p}_g'$ ，整體粒子群們的行為，則是逐漸朝向 $\bar{p}_g'$ 的位置聚集，在這種運作模式類似於大部分隨機搜尋的演算法，也就是集中較大比例的搜尋資源在預期較有可能產生最佳解的區域進行搜尋。以此方式不斷地重複進行以期搜尋到真正最佳解的所在位置。此在方向上引導粒子前進的特性為 PSO 中最基本也是最重要的搜尋精神。

在本文提出的 PSGES 裡，主要將粒子移動方式往 $\bar{p}_g$ 聚集的「概念」與傳統 ES 的突變行為結合，擷取 PSO 裡方向上的引導機制，彌補 ES 中以亂數決定突變方向所造成的不穩定現象。在下一節中，我們將對於引導式突變的運作方式與細部流程加以說明，並比較引導式突變與傳統突變在搜尋行為上的差異。

3.3 引導式突變

在本節中我們將對於本論文所提出的突變機制，稱之為引導式突變做說明。綜合 3.1 與 3.2 節所做的描述，引導式突變為基於關聯式突變架構上，賦予方向引導的另一種突變方式，在突變方向的決定上，模仿了 PSO 中粒

子移動的「概念」，藉由群體合作的資訊來判定預期會有較好的方向，因此在引導式突變架構下，個體突變不再是獨自地以隨機擾動的方式來決定，而是在突變的過程中參考群體智能 (Swarm Intelligence) 所得的資訊，用以促使個體往適合方向前進的機制，與傳統突變的主要差異，為引導式突變在個體演化的行為上，提高機率來突變至預期會有較佳解的區域，個體進行突變的行為將有效降低突變至無效的搜尋空間上。以下將先說明引導式突變的運作機制，接著比較引導式突變與傳統突變兩者之間的差異。

對於一個二維平面上的點，我們可利用下列式子求得兩向量的夾角

$$\alpha = \frac{\vec{p}_g \cdot \vec{a}}{|\vec{p}_g| \times |\vec{a}|} \quad (2.1)$$

其中「 $\cdot$ 」為向量內積， $\vec{a}$ 為目前個體在解空間上的位置， $\vec{p}_g$ 為截至目前為止群體所找到的最佳解， α_g 為個體 $\vec{a}$ 與目前最佳解 $\vec{p}_g$ 之間的向量夾角，其中 $\alpha_g \in [-\pi, \pi]$ 。對於一個 n 維的問題，兩兩座標軸一對，可利用上列公式求得 $n(n-1)/2$ 個旋轉角度，並以向量形式 $\vec{\alpha}_g$ 來表示。與關連性突變不同的是，旋轉角度的決定為在突變的過程中，依據 $\vec{p}_g$ 所在的位置做立即的調整，個體突變的行為能及時依據群體在解空間中的分布行為來進行反應，因此可大幅度降低需被調整的策略參數個數，個體 $\vec{a}$ 的編碼形式可簡化為

$$\vec{a} = (x_1, \dots, x_n; \sigma_1, \dots, \sigma_n) = (\vec{x}; \vec{\sigma})。$$

其中 $x_1, \dots, x_n$ 為目標變數， $\sigma_1, \dots, \sigma_n$ 為每一個座標軸個別的突變強度。

引導式突變的行為可細分成下列幾個步驟：

步驟一：首先針對策略參數 $\bar{\sigma}$ 進行突變運算

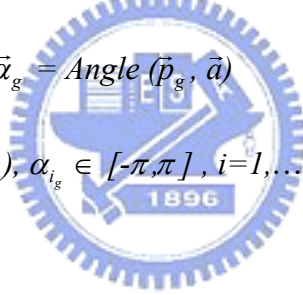
$$\sigma_i^{t+1} = \sigma_i^t \cdot e^{\tau' \cdot N(0,1) + \tau \cdot N_i(0,1)} \quad (2.2)$$

其中 $i \in 1, \dots, n$ ， $\tau' \propto 1/\sqrt{2n}$ 與 $\tau \propto 1/\sqrt{2\sqrt{n}}$ 分別代表全域學習率與區域學習率 [34]。在此步驟中我們保留了傳統演化策略特有的 self-adaptation 來對於突變強度做調整，因此對 step-size 的突變行為與傳統 ES 的運作機制相同。

步驟二：計算個體 $\bar{a}$ 相對於目前最佳解 $\bar{p}_g$ 之間的向量夾角 $\bar{\alpha}_g$

$$\bar{\alpha}_g = \text{Angle}(\bar{p}_g, \bar{a}) \quad (2.3)$$

其中 $\bar{\alpha}_g = (\alpha_{1_g}, \alpha_{2_g}, \dots, \alpha_{n(n-1)/2_g})$, $\alpha_{i_g} \in [-\pi, \pi]$, $i=1, \dots, n(n-1)/2$



步驟三：建立引導式旋轉矩陣 M

$$M = \prod_{p=1}^{n-1} \prod_{q=p+1}^n M_{pq}(\alpha_{j_g}) \quad (2.4)$$

其中 $j = \frac{1}{2}(2n-p)(p+1) - 2n + q$ [9]。旋轉矩陣 $M_{pq}(\alpha_{j_g})$ 中，除了 $m_{pp} = m_{qq} = \cos(\alpha_{j_g})$ ， $m_{pq} = -m_{qp} = \sin(\alpha_{j_g})$ 的元素有值之外，其餘形成一個 $n \times n$ 的單位矩陣。

利用由左至右的矩陣乘法可以有效的建立出引導式關聯矩陣。

步驟四：建立引導式策略向量

$$\vec{G}_s = M \cdot \vec{z} \quad (2.5)$$

其中 $\vec{z} = (z_1, \dots, z_n)$, $z_i \sim \sigma_i \cdot N(0, 1)$ ，為一個平均值為 0，每一個維度以其所對應的 step-sizes 作為標準差所取樣的高斯分佈隨機變數向量。 $M \cdot \vec{z}$ 表示矩陣相乘 ($\vec{G}_s$ 為 $n \times 1$ 的向量)。

步驟五：利用引導式策略向量 $\vec{G}_s$ ，建立引導式突變向量 $\vec{G}_m$

$$\vec{G}_m = \vec{G}_s \cdot \vec{N}(0, 1) \quad (2.6)$$

其中 $\vec{N}(0, 1)$ 代表 $n \times 1$ 的向量，向量每一個維度是以平均值為 0，標準差為 1 所取樣的高斯分佈隨機變數 ($\vec{G}_m$ 為 $n \times 1$ 的向量)。

步驟六：利用引導式突變向量 $\vec{G}_m$ ，並對個體的目标變數 $\vec{x}$ 進行突變運算

$$\vec{x}' = \vec{x} + \vec{G}_m \quad (2.7)$$

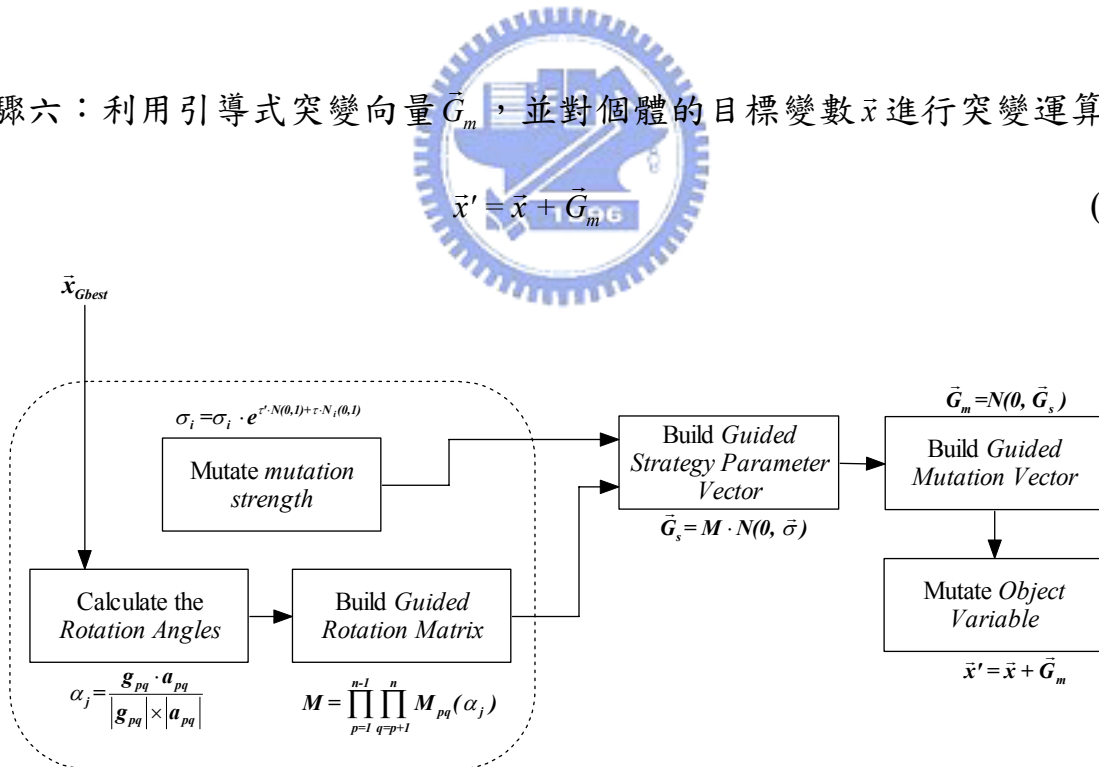


圖 3.2 引導式突變流程圖

經由上列六個步驟，個體的突變橢圓將會轉往 $\vec{p}_g$ 所在的方向，圖 3.2 為引導式突變的流程圖。相對於傳統突變方式以亂數來決定所造成的不穩

定現象，引導式突變在個體突變行為上，模仿 PSO 中粒子逐漸往目前最佳解聚集的特性，以此為概念，來引導個體突變的行為有更高的機率朝向預期會有較佳解的方向前進。圖 3.3 為引導式突變與粒子演算法的示意圖。傳統的突變方式偏向於單一個體在自己周圍隨機的搜尋，而引導式突變引進 PSO 的引導觀念，利用群體智能所獲得的資訊，將在突變的效率上能更有效的突變至預期會有較佳解的方向。圖 3.4 為傳統突變方式與引導式突變的差異比較。

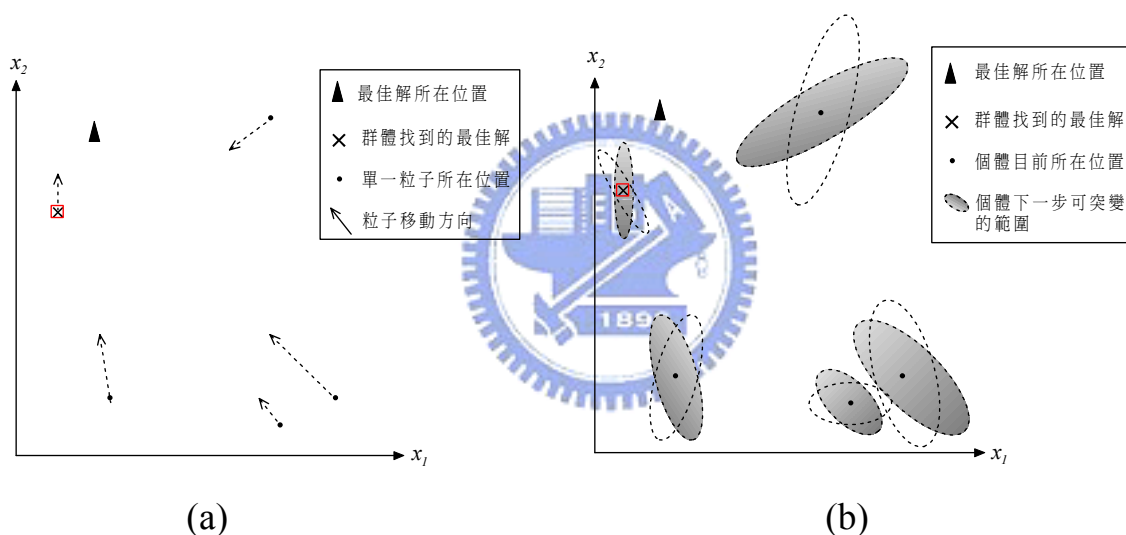
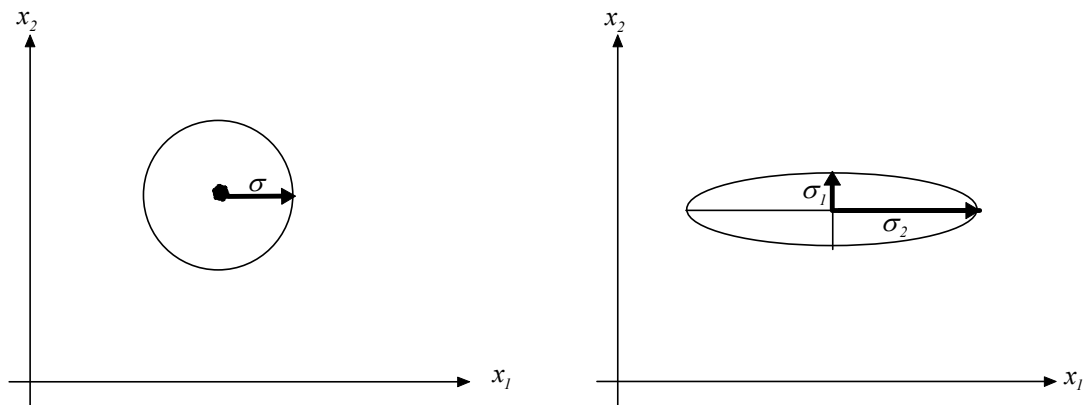
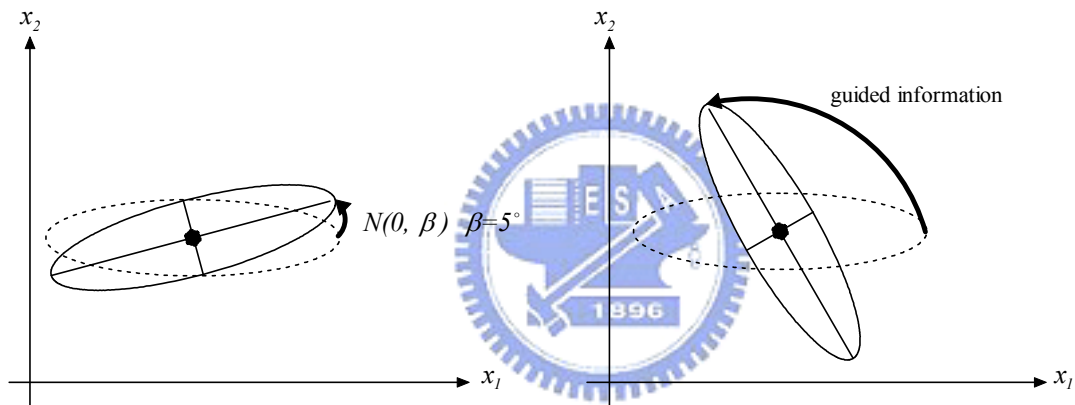


圖 3.3 引導式突變與粒子移動行為比較。(a) 為粒子演算法粒子下一步移動行為示意圖，方點為目前為止所找到的最佳解；(b) 引導式突變決定個體下一步可突變的範圍示意圖。虛線橢圓為傳統的突變橢圓區域，灰色橢圓為引導式突變的突變橢圓區域。



(a) 單一突變強度的非關聯式突變 (b) n 維突變強度的非關聯式突變



(c) 關聯式突變機制

(d) 引導式突變機制

圖 3.4 傳統突變方式與引導式突變比較

在引導式突變中，個體突變橢圓與座標軸獨立的優點仍與以保持，確切地說，我們是以傳統關聯式突變機制作為基礎並加以修改，結合 PSO 的搜尋「概念」進而發展出新的突變方式。我們保留 ES 中最主要的精神，也就是利用 self-adaptation 來對於引導式突變的 step-sizes 做調整；對於方向的

決定上，則是擷取 PSO 中，方向引導的概念機制來做決定。因此，在引導式突變的行為上，為結合兩種實數參數最佳化方法，利用各自領域的主要搜尋精神，所發展出另一種有效的突變機制，在下一節裡，我們將以一個視覺化的角度，來觀察突變運算在被引導的情況下，其突變行為與傳統的方式的比較。

3.4 視覺化評估與觀察

為了有效驗證本文所提的引導式突變在搜尋效能上的可行性，這一節裡，我們以一個視覺化的實驗方式來觀察並比較引導式突變與傳統突變方式之間的差異。我們分成兩個部分來分析，第一個部分為觀察引導式突變相較於 n 維突變強度的非關聯式突變與關連性突變的差異；在實驗的第二個部份，我們觀察放寬旋轉角度限制後，關連性突變其突變橢圓分布行為。

為了觀察上的方便，以及突顯引導式突變在方向決定上與傳統突變方式的不同，在本實驗的設計中，我們將不使用重組的機制，此外，每一個個體利用突變機制產生 80 個子代，篩選機制將 80 個子代中擁有最佳適應值的個體保留。我們以二維的 Sphere Model 作為目標函式，目標函式如公式(2.8)所示，此為一個最小化的問題，最佳解位在 $[0,0]$ 的位置。

$$f(\vec{x}) = \sum_{i=1}^n (x_i)^2, \quad n=2 \quad (2.8)$$

全部實驗皆以相同的初始條件作為起始環境，以此來觀察突變橢圓的分布行為，其初始個體的分布行為如圖 3.4 所示。

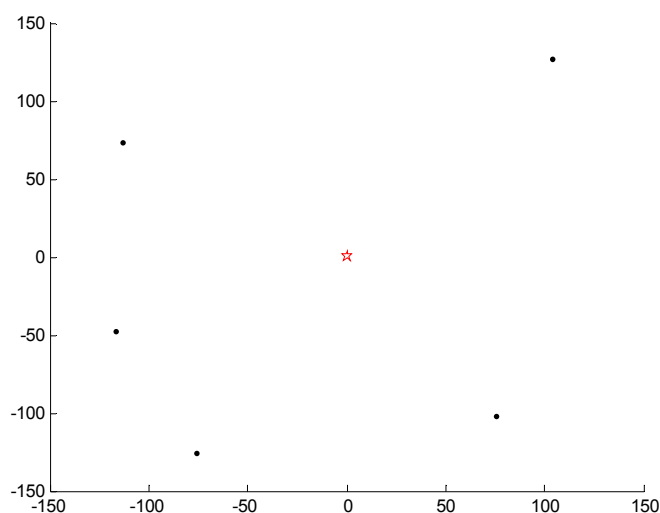


圖 3.5 起始環境下個體的分布情形

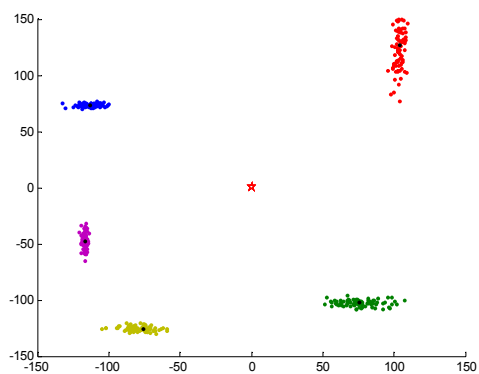
3.4.1 引導式突變與傳統突變方式之比較

此實驗的目的主要著重在探討個體突變後子代的分布行為，比較在引入 PSO 引導機制的概念下，引導式突變與傳統突變方式的差異。圖 3.6、圖 3.7 與圖 3.8 分別為 n 維突變強度的非關聯式突變、關聯式突變與引導式突變的個體突變橢圓分佈情形。

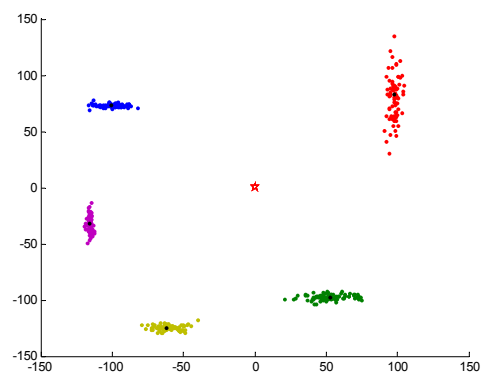
在圖 3.6 的 n 維突變強度的非關聯式突變下，個體所產生的突變橢圓，在突變的方向受限於座標軸系統，因此，只能夠朝向與座標軸相依的方向

做調整，在此突變機制下，主要利用 self-adaptation 來對於 step-sizes 做調整。

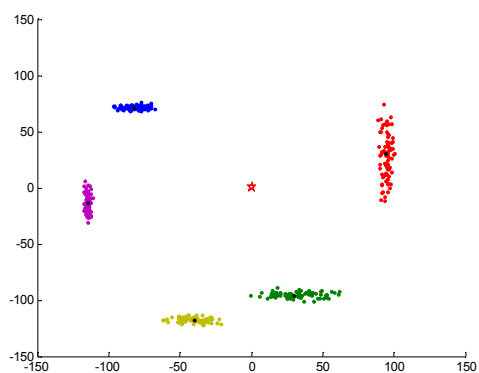
在圖 3.7 的關聯式突變中，我們可觀察到其突變橢圓可達到與座標軸獨立的特性，在方向的決定上，利用 5° 的變異量來對旋轉角度做逐步調整，self-adaptation 雖然在某些情況下會調整 rotation angles 往適合的方向旋轉，然受限於大量旋轉角度需被調整的情況下，篩選機制在策略參數層級造成的高度不穩定性，在大部分的情況仍會造成旋轉錯誤情形，也就是說在突變方向的決定上仍存在著高度不確定性的因子，放大了產生無效的突變行為的機率。圖 3.8 為本文所提的引導式突變的突變橢圓分佈行為，擷取粒子演算法中方向引導的概念，由圖中可觀察出個體突變橢圓分佈，在突變橢圓旋轉的方向依據群體相互合作的資訊，來引導個體往預期會有較佳解的方向，子代的個體將有更高的機率朝向適合的方向進行突變，因此能產生更有效率的突變行為。



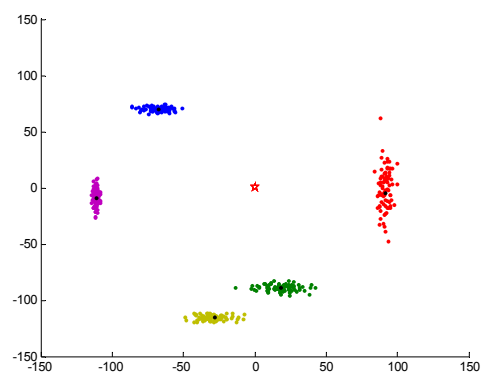
Generation 1



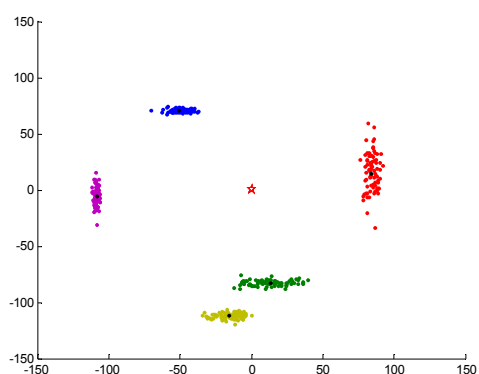
Generation 2



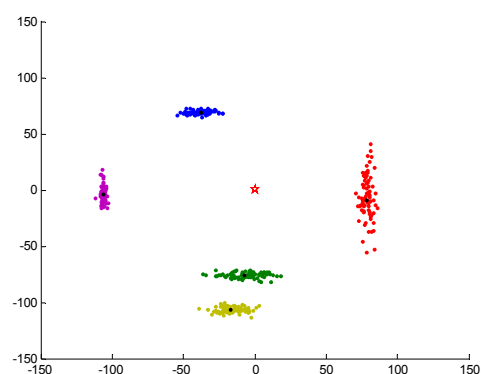
Generation 3



Generation 4

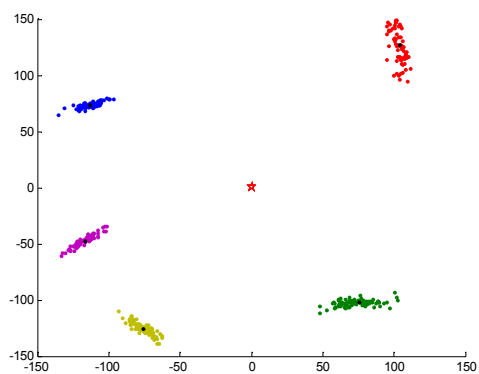


Generation 5

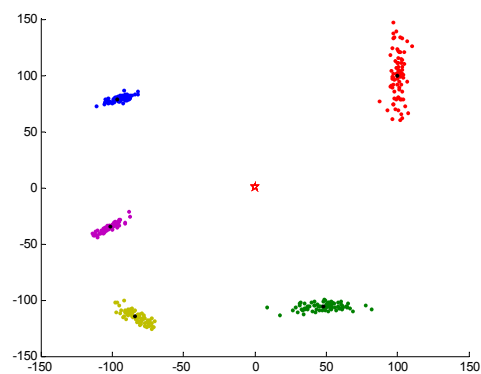


Generation 6

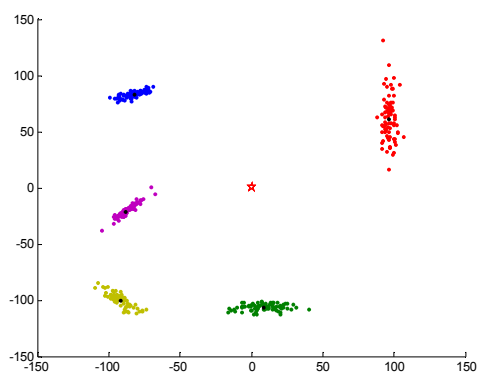
圖 3.6 n 維突變強度的非關聯式突變演化示意圖



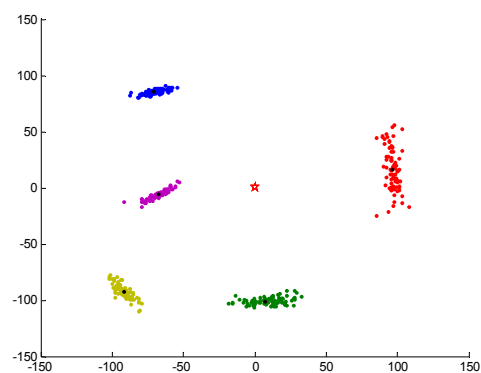
Generation 1



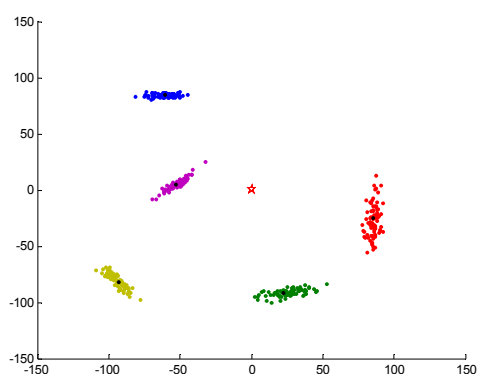
Generation 2



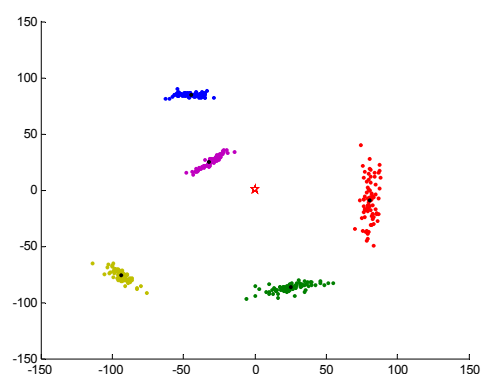
Generation 3



Generation 4

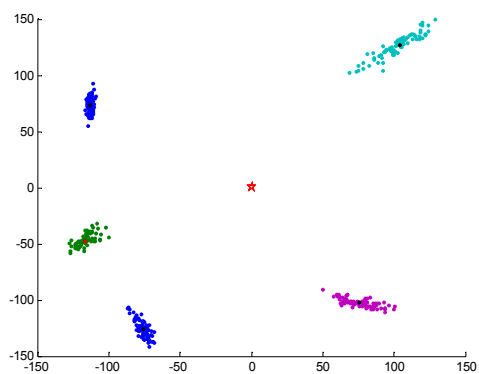


Generation 5

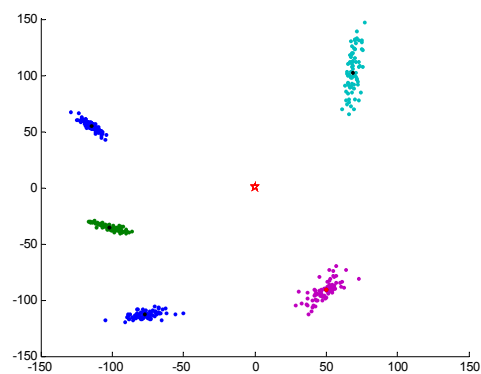


Generation 6

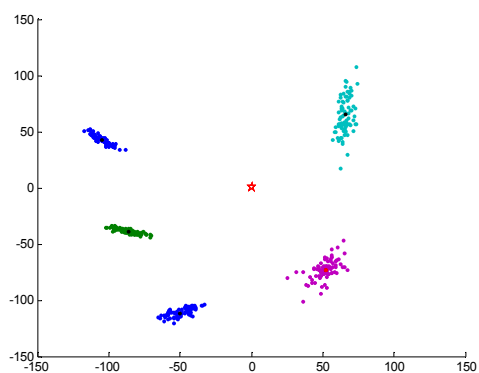
圖 3.7 關聯式突變演化示意圖



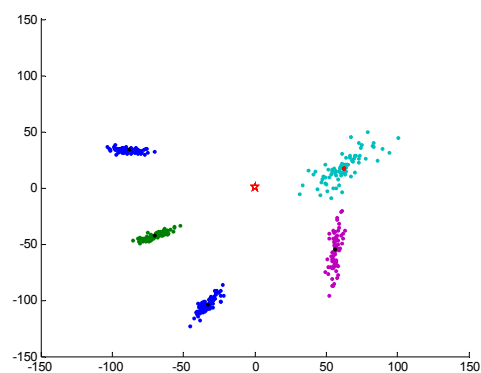
Generation 1



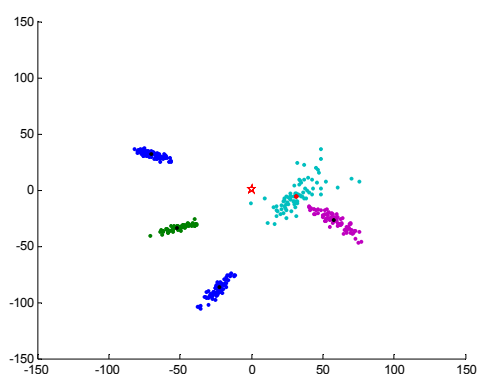
Generation 2



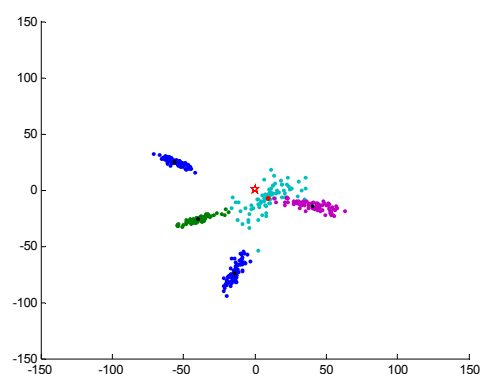
Generation 3



Generation 4



Generation 5



Generation 6

圖 3.8 引導式突變演化示意圖

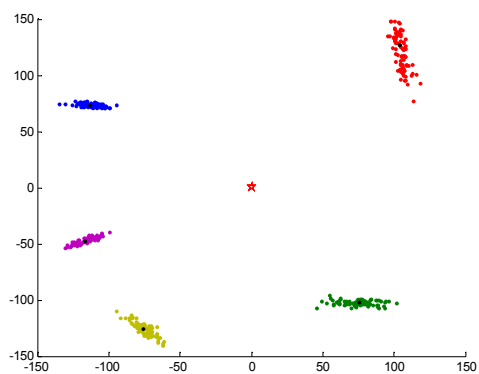
3.4.2 引導式突變與關聯式突變分析

在視覺化觀察的第二部份，我們對於關聯式突變在 rotation angles 的改變量(change rate)做探討。此實驗的觀察重點，在於觀察放大關聯性突變中，rotation angles 的學習率 β 值，對於個體的突變行為所造成的影響。一方面，我們希望在 rotation angles 的改變量 (Change rate)越大越好，其原因是在一次突變的行為中可達到較大的改變，並且以較短的時間來對於 rotation angles 做調整；然而另一方面，起因於篩選機制在策略參數層級的挑選限制，rotation angles 的改變量必須有一個上限，否則對於所挑選的個體將不能依存在有效的資訊上而造成隨機改變的行為 (Stochastic behavior)。在這個實驗中，我們放寬其對於 rotation angles 的限制，我們設定 β 值為 20 度與 360 度，並觀察其演化的行為，圖 3.9 為 $\beta=20^\circ$ 的關聯式突變，圖 3.10 為 $\beta=360^\circ$ 的關連性突變。在 $\beta=20^\circ$ 的情況下，個體在單一一次突變的旋轉角度變異量可有 20° 的改變，在 $\beta=360^\circ$ 的情況下，個體在單一一次突變的旋轉角度變異量可有 360° 的改變，也就是可在一次突變中，即可朝任意方向旋轉。

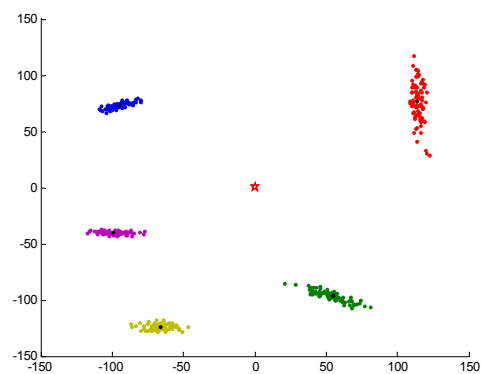
由實驗圖中可發現，即使放寬了單一一次的突變行為對於其旋轉角度上的限制，具有引導式的突變仍然較 $\beta=20^\circ$ 與 $\beta=360^\circ$ 的關聯式突變來的有效率，主要的原因是引導式突變雖然單一一次的突變中放大了旋轉角度

的改變量，但其旋轉角度的決定依存在明確的引導資訊上，因此可同時實現現在旋轉角度的改變量大並以較短的時間來對於旋轉角度做調整，而又不會有在方向上隨機改變的行為產生。

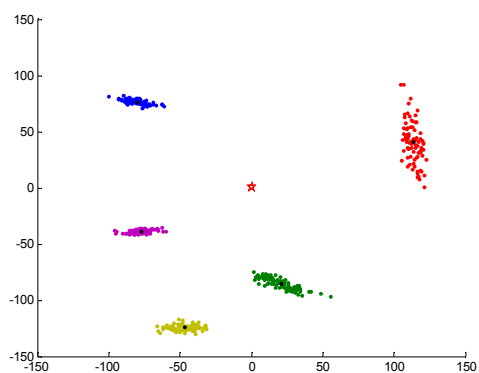




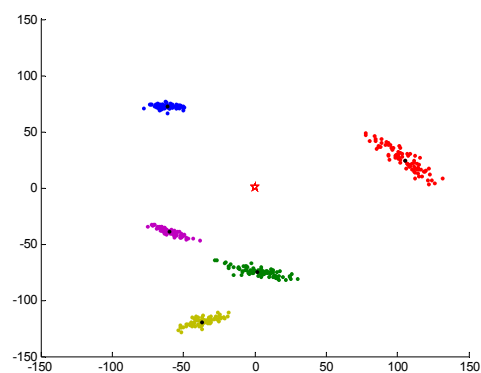
Generation 1



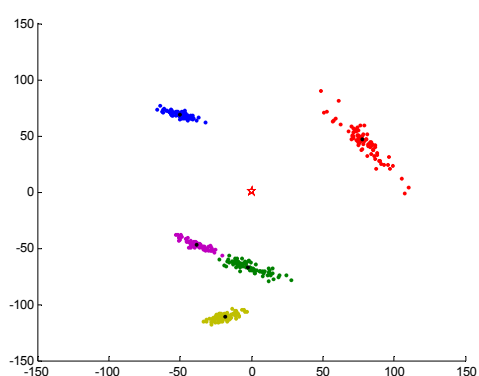
Generation 2



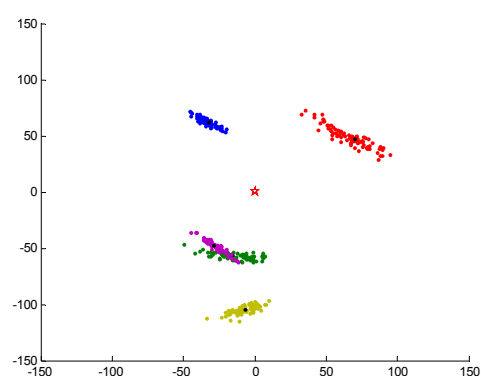
Generation 3



Generation 4

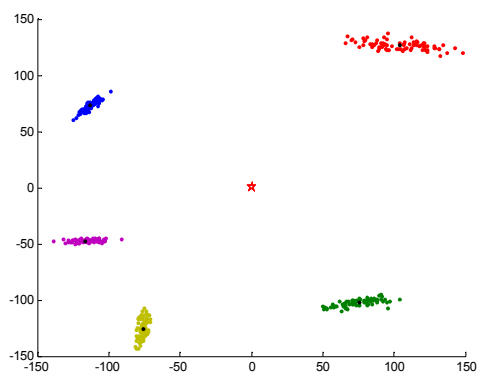


Generation 5

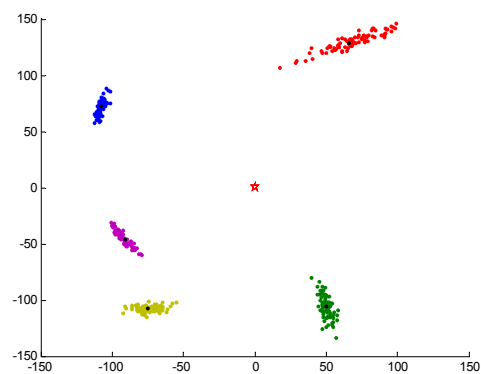


Generation 6

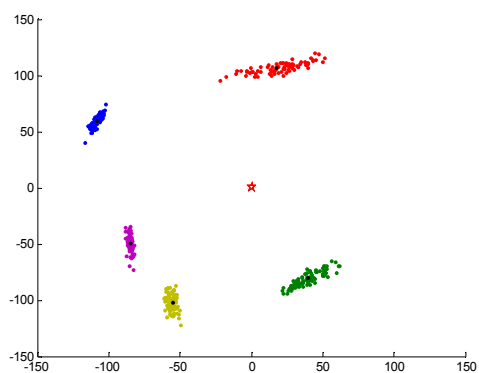
圖 3.9 關聯式突變 ($\beta=20^\circ$) 演化示意圖



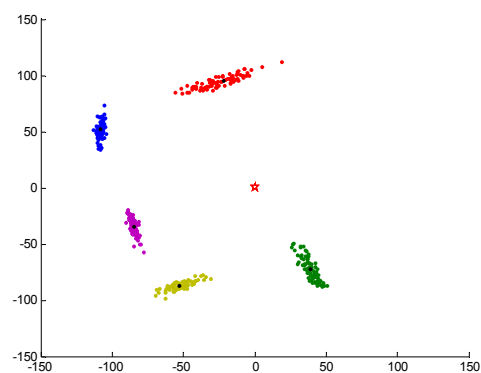
Generation 1



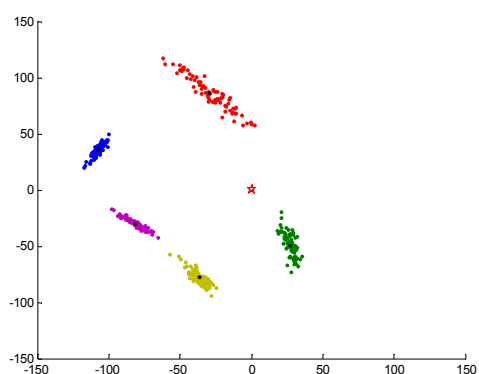
Generation 2



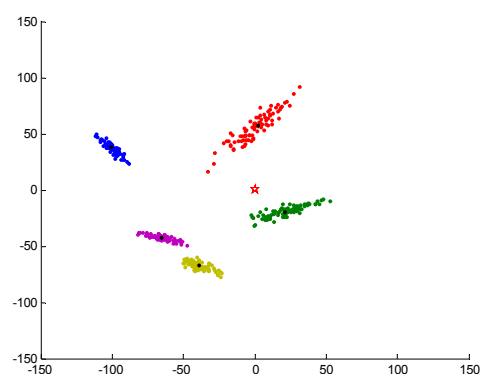
Generation 3



Generation 4



Generation 5



Generation 6

圖 3.10 關聯式突變 ($\beta=360^\circ$) 演化示意圖

3.5 引導式演化策略

本論文的主要重點為藉由 PSO 中的引導觀念，與傳統 ES 做結合，進而期望在求解問題的效率上能達到效能提升的效果。為了達到此目的，我們特別針對突變運算做改良，並設計了引導式突變的機制。PSGES 主要是將傳統 ES 的突變機制以引導式突變來取代，演化的流程有些許的更動，以下將針對 PSGES 的運作機制加以說明。

不同於傳統 ES 單點各自搜尋的方式，PSGES 為一個利用群體合作的觀念來引導整個群體往預期較有可能產生適合解的方向移動，因此一個族群式 (Population Based) 的演化方式為必需的條件。排除了 $(I+I)$ -ES 與 $(I+\lambda)$ -ES 兩種演化方式，我們所採取的是 (μ, λ) -ES 來做為 PSGES 的基本架構。單一個體 $\vec{I}$ 的表示法分成目標變數 $\vec{x}$ 與 策略參數 $\vec{\sigma}$ 兩個部份

$$\vec{I} = (\vec{x}; \vec{\sigma}) = (x_1, \dots, x_n; \sigma_1, \dots, \sigma_n)$$

除此之外，對於旋轉角度 $\vec{\alpha}$ 的部分，由於是在演化的過程依據群體分佈的資訊即時取的，因此在個體的表示法中可予以省略。

在最佳化的程序開始前，我們首先要對父代的群體產生起始化

$$P(0) = \{\vec{a}_1(0), \dots, \vec{a}_\mu(0)\}$$

初始化的方式與傳統 ES 相同，起始點的位置可由使用者設定，對於一些有明顯條件限制的問題來說，起始點的位置可在合理區域範圍(feasible region)

內以 n 維的常態機率分佈 (Uniform Probability Distribution) 隨機取得。給定一個目標函式 (Object Function)，對於父代的個體計算其適應值 (Fitness Value)

$$f(\vec{a}) = \phi(\vec{x}) \quad (2.9)$$

模仿 PSO 的運作模式，我們將擁有最佳適應值的個體 $\vec{g}$ 記錄下來，並以此點作為全域性引導的資訊。接著重複進行下列步驟直到中止條件達成為止



步驟一：利用重組運算產生 λ 個子代個體，。

步驟二：參考 $\vec{p}_g$ 所在位置，對 λ 個子代個體進行引導式突變。

步驟三：計算 λ 個子代個體的適應值並予以排列，取得擁有最佳適應值的個體 $\vec{g}'$ 。

步驟四：若 $\vec{p}_g'$ 優於 $\vec{p}_g$ ，則將 $\vec{p}_g$ 替換成 $\vec{p}_g'$ ，否則 $\vec{p}_g$ 予以保留。

步驟五：利用篩選機制挑選擁有最佳適應值的 μ 個個體做為下一代的父代人口群。

步驟六：若符合終止條件則停止最佳化過程，否則重複步驟一到五。

圖 3.11 為 PSGES 的概念示意圖，在 PSGES 的運作概念上，傳統 ES

所扮演的角色較偏向於單點的隨機區域搜尋，PSO 全域性最佳解引導的特性，在 PSGES 中則扮演著偏向全域性搜尋的角色。結合了傳統 ES 中偏向於區域搜尋的機制與 PSO 裡整體最佳解的全域性搜尋機制，在 PSGES 的演化方式中，我們直接藉由將個體的突變橢圓轉向全域性最佳解的方向來達成。因此，在決定個體 step-sizes 部份，利用傳統 ES 中的 self-adaptation 的機制，將能有效的調節突變的強度並有利於區域搜尋的效能提升；在個體突變的方向部份，個體所突變後的子代，相較於傳統 ES 來說，將有更高的機率朝向預期會有較佳解的方向前進，此舉將有利於全域性的搜尋。整體來說，對於個體突變的行為上，不再是單點隨機搜尋，而是參與了群體合作的資訊，更近一步說，我們直接將 PSO 引導的觀念，利用在突變方向上的引導來達成。圖 3.12 為引導式突變的流程圖。



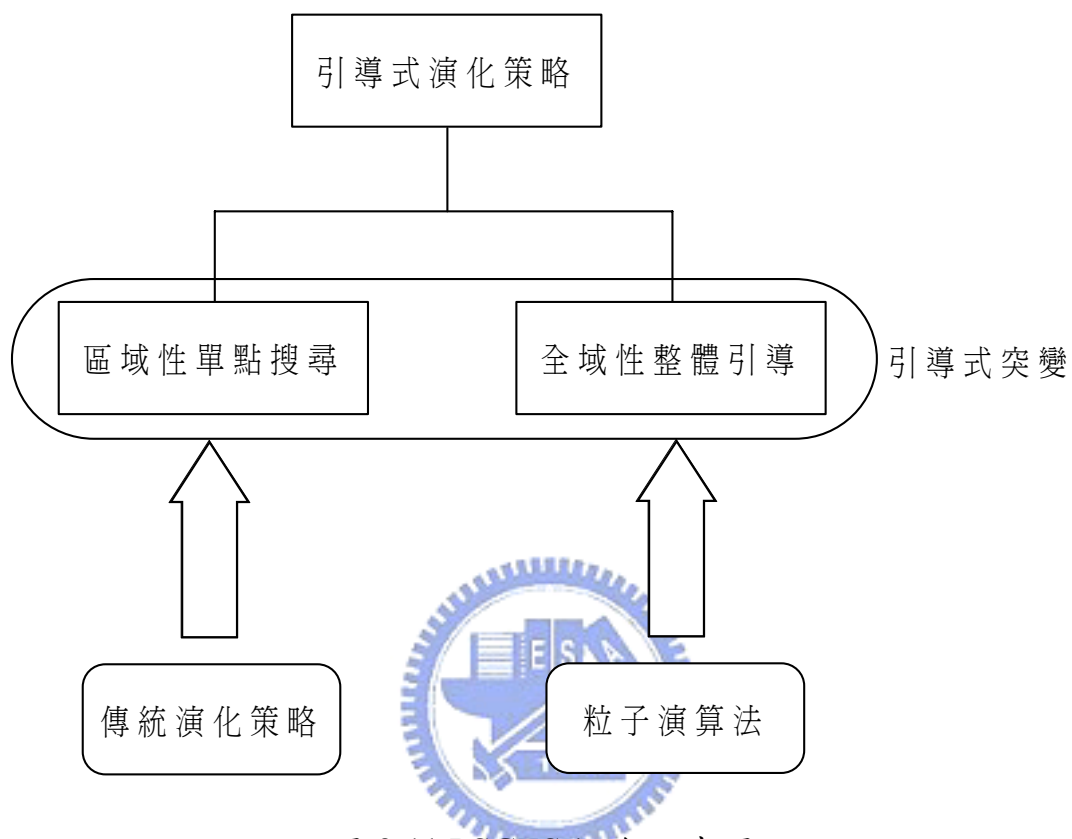


圖 3.11 PSGES 概念示意圖

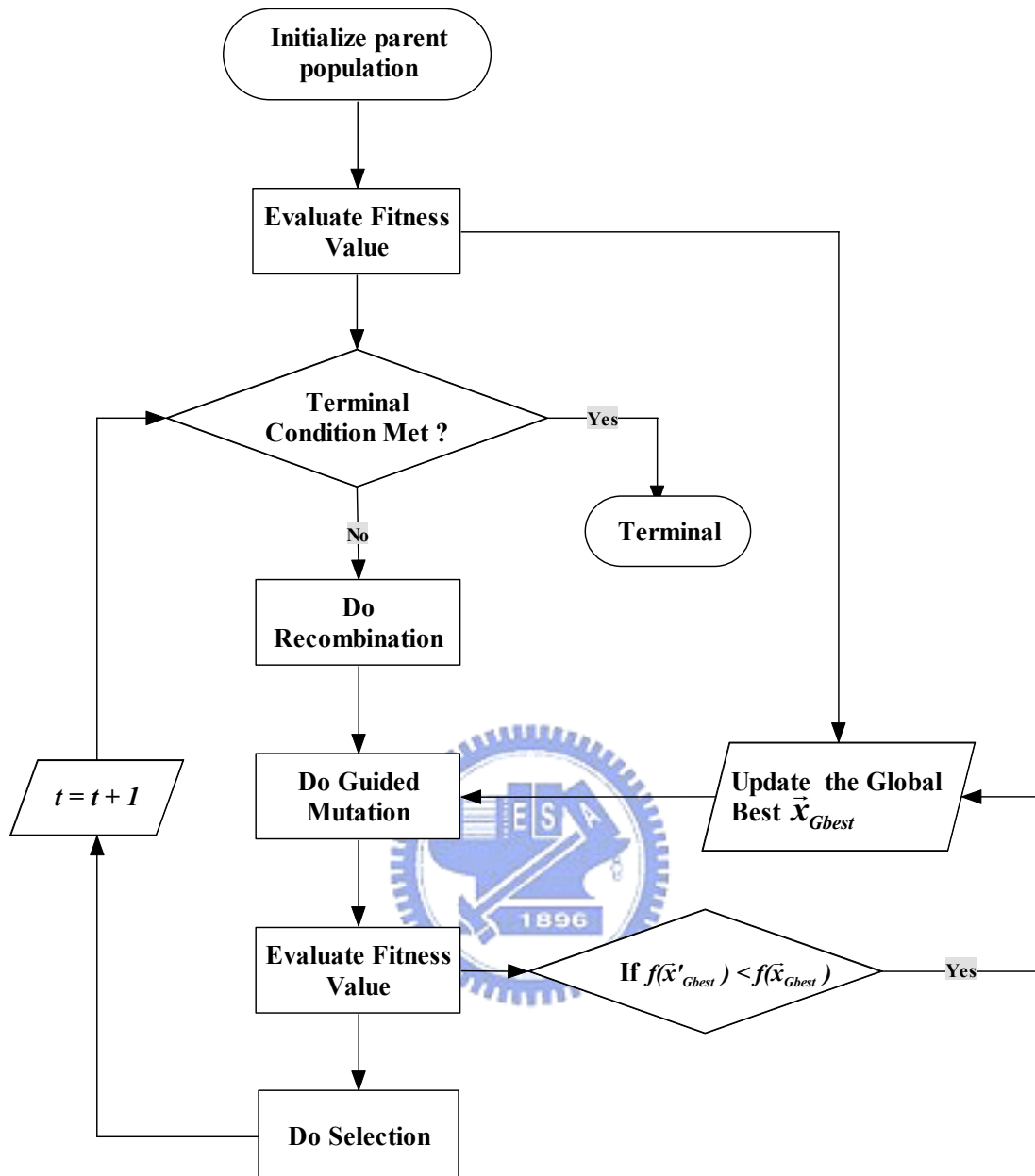


圖 3.12 引導式演化策略流程圖

3.6 本章總結

在本章中，我們首先介紹了 PSGES 的想法起源，觀察傳統 ES 的運作機制，發現個體在突變方向的決定上較缺乏一個決定性的因素來促使群體往適合的方向前進，因此我們藉由 PSO 中，粒子們相互引導的機制，結合至傳統 ES 的突變行為上，我們首先描述了如何基於關聯式突變並加以修改，發展了結合 PSO 於 rotation angles 在方向上的引導與 ES 中 step-sizes 以 self-adaptation 對其做調整的機制，試圖在全域性搜尋與區域性搜尋上達到良好的平衡，最後我們描述了 PSGES 的完整架構，並對於其優缺點及相關特性做討論。



第四章 實驗結果與效能分析

在本文所提出的 PSGES，為針對實數參數最佳化問題所發展的一套演算法，為了公平及有效地衡量 PSGES 在求解問題上的效能。我們以 CEC 2005 special session (IEEE Congress on Evolutionary Computation of the Special Session on Real-Parameter Optimization) 在 2005 年所提供的標準測試函式來做為測試的平台，基於這個測試平台的評判標準下，目前已有釋出在此評判標準下所得的相關結果，我們的比較對象將與在這個 session 已 accepted 的結果進行比較與分析，進而觀察引導式演化策略於搜尋效能上所提升的程度。

在第一節我們將解釋與說明為何挑選此函式作為實驗衡量的標準；在第二節裡，我們將對於所要處理的測試函式針對其特性作說明；第三節為關於實驗參數設計的部份；第四節為基於這組 benchmark 下的實驗結果與比較；並以第五節為本章的總結。

4.1 實驗衡量標準

演化計算領域近三十年的發展過程中，至今已有許多不同型式的演算法被設計出來以求解實數參數的最佳化問題，一些較廣為人知的方法包括

實數參數的遺傳演算法 (Real-Parameter EAs)、演化式規劃 (Evolutionary Programming, EP)、演化策略 (Evolution Strategies, ES)、粒子演算法 (Particle Swarm Optimization, PSO)、差值演化方法 (Differential Evolution, DE)，一些較古典的方法如修正牛頓法 (quasi-Newton method, QN)，古典雜交式演化方法 (hybrid evolutionary-classical methods)，以及其他非演化式的方法如模擬降溫法 (simulated annealing, SA)、禁區搜尋法 (tabu search, TS)等。在這些最佳化的演算法中，許多被提出的方法在運算元以及其運作機制上有著明顯不同的差異，在大部分的比較方式，通常是考慮一些較常使用的測試函式如 Sphere, Schwefel's, Rosenbrock's, Rastrigin's 等以及其相關的變形。雖然說一些方法的比較結果可在文獻中得知，然而對於所挑選的測試函式往往容易使人混淆，在某些情況下，所提的演算法本身與測試函式之間關係完全是人為操縱與挑選，因此可能發生的情況是所提出來的方法只在某些特定的測試函式下才能運作。為了公平地比較演算法之間的效能差異，本文利用CEC (IEEE Congress on Evolutionary Computation)在2005年所提供的一系列標準測試函式作為測試的平台[37]，在衡量演算法的效能上使用更有系統性的方式來做評判，藉由明確地定義出條件的設定，譬如終止條件的判斷、人口群的大小設定、起始條件的限制等。基於相同環境設定下所做的比較結果，我們將可更客觀地對於演算法的效能上做分析與探討。在下一節中，我們將對於所要測試的標準測試函式作一完整的描述。

4.2 測試函式

此組 benchmark 總共包含了 25 個 test functions，依其性質可區分成兩大部分，其中 5 個為 unimodal functions，20 個為 multimodal functions，multimodal functions 依其組成又可區分成三個部份：basic functions、expanded functions 與 hybrid composition functions。對此 25 個測試函式的整理如表 4.1 所示，詳細的測試函式定義放置於附錄一。

此組 benchmark 皆為實數參數最小化的問題，每一個測試函式皆有其不同的特性，真正的最佳解 $f(x^*)$ 做過 *bias* 的處理。此 25 個函式在解空間的搜尋範圍除了 Function 7 與 Function 25 沒有限制之外，其餘皆有固定的搜尋範圍，起始範圍的設定皆有明確的定義並整理於表 4.2。實驗數據的取得為針對每一個函式執行 100000 次函式計算(Function Evaluations, FES)，並重複執行 25 次實驗。在演化次數(FES)達到 1000、10000、100000 時，分別記錄錯誤值(error value) $f(x) - f(x^*)$ 中 1st(Best)、7th、13th(Median)、19th、25th(Worst)所得到的值，以及計算每一個 function 在 25 次的實驗所得的平均值與標準差。當終止條件達成時，依照問題的困難度分別定義一個已求解之臨界值 (Solved Threshold)，若最佳化的結果其錯誤率小於所定義的臨界值 $f(x) - f(x^*) \leq Threshod$ ，則我們視此問題為可被解決的問題。表 4.3 為 25 個 test functions 的求解臨界值定義。

表 4.1 25 個 test functions 特性分類

Unimodal Functions (5):	
□ Basic Functions (5)	
F_1 :	Shifted Sphere Function
F_2 :	Shifted Schwefel's Problem 1.2
F_3 :	Shifted Rotated High Conditioned Elliptic Function
F_4 :	Shifted Schwefel's Problem 1.2 with Noise in Fitness
F_5 :	Schwefel's Problem 2.6 with Global Optimum on Bounds
Multimodal Functions (20):	
□ Basic Functions (7):	
F_6 :	Shifted Rosenbrock's Function
F_7 :	Shifted Rotated Griewank's Function without Bounds
F_8 :	Shifted Rotated Ackley's Function with Global Optimum on Bounds
F_9 :	Shifted Rastrigin's Function
F_{10} :	Shifted Rotated Rastrigin's Function
F_{11} :	Shifted Rotated Weierstrass Function
F_{12} :	Schwefel's Problem 2.13
□ Expanded Functions (2):	
F_{13} :	Expanded Extended Griewank's plus Rosenbrock's Function (F8F2)
F_{14} :	Shifted Rotated Expanded Scaffer's F6
□ Hybrid Composition Functions (11):	
F_{15} :	Hybrid Composition Function
F_{16} :	Rotated Hybrid Composition Function
F_{17} :	Rotated Hybrid Composition Function with Noise in Fitness
F_{18} :	Rotated Hybrid Composition Function
F_{19} :	Rotated Hybrid Composition Function with a Narrow Basin for the Global Optimum
F_{20} :	Rotated Hybrid Composition Function with the Global Optimum on the Bounds
F_{21} :	Rotated Hybrid Composition Function
F_{22} :	Rotated Hybrid Composition Function with High Condition Number Matrix
F_{23} :	Non-Continuous Rotated Hybrid Composition Function
F_{24} :	Rotated Hybrid Composition Function
F_{25} :	Rotated Hybrid Composition Function without Bounds

表 4.2 最佳解、搜尋範圍與起始範圍定義表

Function Number	$f(x^*)$	Search Range	Initialization Range
$F1$	-450	$[-100,100]$	$[-100,100]$
$F2$	-450	$[-100,100]$	$[-100,100]$
$F3$	-450	$[-100,100]$	$[-100,100]$
$F4$	-450	$[-100,100]$	$[-100,100]$
$F5$	-310	$[-100,100]$	$[-100,100]$
$F6$	390	$[-100,100]$	$[-100,100]$
$F7$	-180	$[-\infty,\infty]$	$[0,600]$
$F8$	-140	$[-32,32]$	$[-32,32]$
$F9$	-330	$[-5,5]$	$[-5,5]$
$F10$	-330	$[-5,5]$	$[-5,5]$
$F11$	90	$[-0.5,0.5]$	$[-0.5,0.5]$
$F12$	-460	$[-\pi,\pi]$	$[-\pi,\pi]$
$F13$	-130	$[-5,5]$	$[-5,5]$
$F14$	-300	$[-100,100]$	$[-100,100]$
$F15$	120	$[-5,5]$	$[-5,5]$
$F16$	120	$[-5,5]$	$[-5,5]$
$F17$	120	$[-5,5]$	$[-5,5]$
$F18$	10	$[-5,5]$	$[-5,5]$
$F19$	10	$[-5,5]$	$[-5,5]$
$F20$	10	$[-5,5]$	$[-5,5]$
$F21$	360	$[-5,5]$	$[-5,5]$
$F22$	360	$[-5,5]$	$[-5,5]$
$F23$	360	$[-5,5]$	$[-5,5]$
$F24$	260	$[-5,5]$	$[-5,5]$
$F25$	260	$[-\infty,\infty]$	$[2,5]$

表 4.3 25 個 test functions 之求解臨界值定義

Function	Accuracy	Function	Accuracy
<i>F1</i>	1.00E-6	<i>F14</i>	1.00E-2
<i>F2</i>	1.00E-6	<i>F15</i>	1.00E-2
<i>F3</i>	1.00E-6	<i>F16</i>	1.00E-2
<i>F4</i>	1.00E-6	<i>F17</i>	1.00E-1
<i>F5</i>	1.00E-6	<i>F18</i>	1.00E-1
<i>F6</i>	1.00E-2	<i>F19</i>	1.00E-1
<i>F7</i>	1.00E-2	<i>F20</i>	1.00E-1
<i>F8</i>	1.00E-2	<i>F21</i>	1.00E-1
<i>F9</i>	1.00E-2	<i>F22</i>	1.00E-1
<i>F10</i>	1.00E-2	<i>F23</i>	1.00E-1
<i>F11</i>	1.00E-2	<i>F24</i>	1.00E-1
<i>F12</i>	1.00E-2	<i>F25</i>	1.00E-1
<i>F13</i>	1.00E-2		



4.3 實驗評比對象及參數設定

本文所提的方法為基於傳統 ES 在突變機制上做改良，因此比較的對象分成兩個部份討論。首先我們比較傳統 ES 與 PSGES 在搜尋效能上的差異。此部分的觀察重點，著重在比較改良後的 PSGES，其效能相較於傳統 ES 所提升的程度。對於傳統 ES 在這組 benchmark 下的效能表現，在 CEC05 Special Session 所提供的相關文獻中¹，Lino Costa 於所提的 PLES[38] (A Parameter-less Evolution Strategy for Global Optimization)裡已有最新且完整的實驗數據與結果，因此我們將不自行實作，此外，由於 PLES 也是基於傳統 ES 的架構下，對於其參數調整方面做改良，所以也將之列入我們第一部分的比較對象。第二部份為 PSGES 與其他進階演算法的比較，此部分的觀察重點，著重在觀察 PSGES 相對於其他進階演算法對於求解能力上的比較，基於這組 benchmark 25 個 test functions 下，探討可解決的問題個數。

本實驗參數的設定部份與 PLES 所設定的條件相同，父代人口數(μ)為 10，子代人口數(λ)為 100，參與重組的候選個數(ρ)為 10，突變強度的全域學習率與區域學習率分別為 $\tau' \propto 1/\sqrt{2n}$ 與 $\tau \propto 1/\sqrt{2\sqrt{n}}$ 。在相同的參數設定之下，觀察其結果並與之比較。

¹ http://www.ntu.edu.sg/home/epnsugan/index_files/CEC-05/CEC05.htm

表 4.4 實驗參數設定

Parameter	Value
Dimensions	10
Parent Population Size (μ)	10
Offspring Population Size (λ)	100
Number of Recombinants (ρ)	10
Global Learning Rate (τ)	$1/\sqrt{2n}$
Individual Learning Rate (τ')	$1/\sqrt{2\sqrt{n}}$

4.4 實驗結果與效能分析

關於 PSGES 的實驗結果我們分成兩個部份呈現，第一部份為 PSGES 與傳統 ES 的比較；第二部份為 PSGES 與其他進階演算法的比較。

4.4.1 引導式演化策略與傳統演化策略比較

本文所提的 PSGES 基於這組 benchmark 下的完整實驗數據整理於表 4.5 至表 4.9。表 4.10 為第一部分比較結果，此部份著重在觀察 25 個 test functions 於 25 次獨立實驗下的平均效能表現，觀察 PSGES 效能提升的程度，在相同的參數設定下(如表 4.4)，由實驗結果可知，在五個 unimodal 的 test functions 下，PSGES 的表現皆明顯優於傳統 ES 與 PLES。在 multimodal test

functions 的結果部份，傳統 ES 在 F8、F11、F14 與 F25 有不錯的表現，PLES 於 F16、F17、F24 下有不錯的表現，剩餘 13 個 multimodal 的 test functions，PSGES 的表現皆比傳統 ES 與 PLES 優良。整體來說，在引入群體智能概念下的 PSGES，基於這組 benchmark 下所衡量的結果，對於其效能確實能達到提升的作用。



表 4.5 Result of Function 1-5

Prob. FES		1	2	3	4	5
1e3	1 st	1.1664e+002	2.8102e+003	7.3748e+006	2.1957e+003	1.8381e+003
	7 th	7.5376e+002	3.7987e+003	1.5956e+007	4.3276e+003	5.2436e+003
	13 th	1.1690e+003	4.8008e+003	3.2720e+007	8.3719e+003	5.5585e+003
	19 th	1.7258e+003	6.3490e+003	4.7187e+007	1.3432e+004	7.2428e+003
	25 th	6.0549e+003	1.3680e+004	2.7429e+008	2.7293e+004	9.2892e+003
	Mean	1.5497e+003	5.5578e+003	5.4321e+007	1.0195e+004	5.9378e+003
	Std	1.4673e+003	2.6159e+003	6.6220e+007	6.8490e+003	1.5924e+003
1e4	1 st	5.0129e-009	3.0995e-001	1.3471e+005	3.7569e-001	2.7188e+001
	7 th	1.2691e-008	2.1921e+000	5.8265e+005	5.0152e+000	5.8214e+001
	13 th	7.7210e-008	5.8415e+000	1.4316e+006	1.1721e+001	1.1101e+002
	19 th	2.6070e-007	1.6760e+001	5.1381e+006	2.8441e+001	2.9575e+002
	25 th	2.1939e-006	9.8240e+001	3.7279e+007	7.9032e+002	3.9964e+002
	Mean	2.6641e-007	1.7360e+001	4.8032e+006	5.7165e+001	1.6999e+002
	Std	4.7612e-007	2.4932e+001	8.1569e+006	1.5698e+002	1.2902e+002
1e5	1 st	0.0000e+000	0.0000e+000	2.7709e-004	0.0000e+000	3.8487e+000
	7 th	0.0000e+000	0.0000e+000	3.6220e-004	0.0000e+000	1.7192e+001
	13 th	0.0000e+000	0.0000e+000	4.3326e-004	0.0000e+000	5.3304e+001
	19 th	0.0000e+000	0.0000e+000	4.7238e-004	0.0000e+000	2.0930e+002
	25 th	0.0000e+000	0.0000e+000	7.6356e+001	5.6843e-014	3.0360e+002
	Mean	0.0000e+000	0.0000e+000	3.1711e+000	1.3642e-014	1.0535e+002
	Std	0.0000e+000	0.0000e+000	1.5256e+001	2.4777e-014	9.8162e+001

表 4.6 Result of Function 6-10

Prob. FES		6	7	8	9	10
1e3	1 st	6.1432e+004	1.4548e+003	2.1215e+001	7.1701e+001	7.3489e+001
	7 th	1.2216e+007	1.6911e+003	2.1513e+001	1.0830e+002	1.1542e+002
	13 th	2.6775e+007	1.8651e+003	2.1625e+001	1.2102e+002	1.3737e+002
	19 th	1.7184e+008	1.9441e+003	2.1789e+001	1.5044e+002	1.6428e+002
	25 th	1.6248e+009	2.1517e+003	2.1982e+001	1.8875e+002	3.1686e+002
	Mean	1.4638e+008	1.8163e+003	2.1630e+001	1.2934e+002	1.4577e+002
	Std	3.2867e+008	1.9077e+002	2.0013e-001	2.9149e+001	5.2458e+001
1e4	1 st	2.1162e-001	3.3681e-001	2.0869e+001	4.3911e-006	4.9725e+001
	7 th	6.9529e+000	1.0295e+000	2.1445e+001	2.9849e+000	1.1028e+002
	13 th	5.2195e+001	1.0358e+000	2.1627e+001	3.9798e+000	1.3919e+002
	19 th	1.4878e+003	1.0579e+000	2.1687e+001	4.9755e+000	1.6384e+002
	25 th	1.0937e+004	1.1757e+000	2.2089e+001	9.6389e+001	2.3215e+002
	Mean	1.5646e+003	1.0002e+000	2.1590e+001	1.1844e+001	1.4029e+002
	Std	2.8853e+003	1.8077e-001	2.7337e-001	2.4542e+001	4.3938e+001
1e5	1 st	1.6859e-009	0.0000e+000	2.0429e+001	0.0000e+000	4.9748e+000
	7 th	3.6909e-009	0.0000e+000	2.0844e+001	1.9899e+000	1.0945e+001
	13 th	4.4096e-009	7.3960e-003	2.0949e+001	3.9798e+000	1.2934e+001
	19 th	4.5436e-009	9.8573e-003	2.1001e+001	3.9798e+000	1.8904e+001
	25 th	3.9866e+000	3.2021e-002	2.1223e+001	8.9546e+000	2.9849e+001
	Mean	1.5946e-001	7.3913e-003	2.0909e+001	3.4625e+000	1.4686e+001
	Std	7.9732e-001	7.8216e-003	1.6900e-001	1.9506e+000	6.3402e+000

表 4.7 Result of Function 11-15

Prob. FES		11	12	13	14	15
1e3	1 st	1.6155e+001	4.2662e+004	5.9581e+000	4.3898e+000	6.1281e+002
	7 th	1.7472e+001	8.1683e+004	1.2785e+001	4.7545e+000	6.6278e+002
	13 th	1.8718e+001	1.1919e+005	1.9967e+001	5.0252e+000	7.0737e+002
	19 th	1.9742e+001	1.7524e+005	5.3441e+001	5.1401e+000	8.8134e+002
	25 th	2.4505e+001	4.1103e+005	2.7918e+003	5.7078e+000	1.1801e+003
	Mean	1.8859e+001	1.4501e+005	1.7175e+002	4.9931e+000	7.9613e+002
	Std	1.9822e+000	9.4094e+004	5.5781e+002	2.8902e-001	1.8703e+002
1e4	1 st	1.7480e+001	2.0385e-001	7.6817e+000	4.7376e+000	7.8552e+001
	7 th	1.9429e+001	2.0778e+001	9.2818e+000	4.9999e+000	3.0028e+002
	13 th	2.0289e+001	1.1756e+002	1.1243e+001	5.0000e+000	4.0000e+002
	19 th	2.1952e+001	7.2173e+002	1.3052e+001	5.0003e+000	4.0626e+002
	25 th	2.5406e+001	2.7261e+003	1.1143e+002	5.4875e+000	5.0000e+002
	Mean	2.0684e+001	5.0224e+002	1.8576e+001	5.0249e+000	3.2660e+002
	Std	2.1435e+000	7.3999e+002	2.4479e+001	1.3249e-001	1.2989e+002
1e5	1 st	1.0853e+001	7.0949e-008	0.0000e+000	5.0000e+000	7.8515e+001
	7 th	1.3019e+001	9.3505e-007	5.0999e-001	5.0000e+000	3.0000e+002
	13 th	1.3608e+001	1.0003e+001	7.3419e-001	5.0000e+000	4.0000e+002
	19 th	1.4574e+001	7.1225e+002	1.0738e+000	5.0000e+000	4.0626e+002
	25 th	1.5038e+001	1.6936e+003	3.2384e+000	5.0000e+000	5.0000e+002
	Mean	1.3546e+001	3.6095e+002	8.2182e-001	5.0000e+000	3.2644e+002
	Std	1.1995e+000	6.1590e+002	6.7026e-001	0.0000e+000	1.2996e+002

表 4.8 Result of Function 16-20

Prob. FES		16	17	18	19	20
1e3	1 st	2.9763e+002	3.6516e+002	1.0097e+003	9.8585e+002	9.4818e+002
	7 th	3.9090e+002	4.1410e+002	1.1689e+003	1.1464e+003	1.2005e+003
	13 th	4.6418e+002	4.8908e+002	1.1953e+003	1.2061e+003	1.2372e+003
	19 th	5.4648e+002	5.8575e+002	1.2901e+003	1.2958e+003	1.2958e+003
	25 th	6.2830e+002	9.0146e+002	1.6529e+003	1.7951e+003	2.4465e+003
	Mean	4.6606e+002	5.3377e+002	1.2447e+003	1.2524e+003	1.3161e+003
	Std	9.5544e+001	1.4696e+002	1.4729e+002	1.8580e+002	3.0453e+002
1e4	1 st	2.8548e+002	2.7988e+002	3.0000e+002	3.0000e+002	3.0000e+002
	7 th	3.5822e+002	3.9810e+002	8.0019e+002	3.1180e+002	8.0003e+002
	13 th	4.1700e+002	5.0775e+002	9.1338e+002	8.0045e+002	8.0054e+002
	19 th	5.3137e+002	5.4191e+002	9.8192e+002	9.4535e+002	8.7123e+002
	25 th	8.0491e+002	1.0361e+003	1.2787e+003	1.3228e+003	1.2418e+003
	Mean	4.6522e+002	5.0962e+002	8.4075e+002	7.2279e+002	7.5353e+002
	Std	1.3716e+002	1.5508e+002	3.1111e+002	3.1730e+002	2.8141e+002
1e5	1 st	1.1979e+002	1.8356e+002	3.0000e+002	3.0000e+002	3.0000e+002
	7 th	1.5218e+002	2.6777e+002	3.0000e+002	3.0000e+002	8.0000e+002
	13 th	2.1773e+002	2.9307e+002	8.0000e+002	8.0000e+002	8.0000e+002
	19 th	2.3850e+002	3.5288e+002	9.4686e+002	8.2144e+002	8.2132e+002
	25 th	2.8305e+002	4.6059e+002	1.0033e+003	1.0119e+003	1.0024e+003
	Mean	2.0136e+002	3.0325e+002	7.1587e+002	6.6960e+002	7.0579e+002
	Std	4.8440e+001	5.6720e+001	2.7344e+002	2.6511e+002	2.3897e+002

表 4.9 Result of Function 21-25

Prob. FES		21	22	23	24	25
1e3	1 st	1.1304e+003	9.5099e+002	1.0436e+003	5.3267e+002	5.8378e+002
	7 th	1.3093e+003	1.0191e+003	1.3126e+003	8.1422e+002	9.0327e+002
	13 th	1.3365e+003	1.0841e+003	1.3430e+003	9.6853e+002	1.1047e+003
	19 th	1.3681e+003	1.1733e+003	1.3786e+003	1.2063e+003	1.3009e+003
	25 th	1.5673e+003	3.9441e+003	1.4580e+003	2.7783e+003	3.0261e+003
	Mean	1.3327e+003	1.2838e+003	1.3323e+003	1.0673e+003	1.1512e+003
	Std	8.2372e+001	6.2417e+002	9.2305e+001	4.3473e+002	4.6227e+002
1e4	1 st	3.0000e+002	8.1245e+002	5.5947e+002	2.0000e+002	4.0332e+002
	7 th	8.0000e+002	8.4046e+002	9.7050e+002	4.2596e+002	4.2735e+002
	13 th	8.0001e+002	8.6579e+002	1.2423e+003	4.4847e+002	4.4635e+002
	19 th	1.1641e+003	9.1706e+002	1.2671e+003	5.0000e+002	4.7801e+002
	25 th	1.2502e+003	1.0581e+003	1.2950e+003	2.3091e+003	8.6068e+002
	Mean	8.9067e+002	8.8382e+002	1.1395e+003	5.2987e+002	4.7089e+002
	Std	3.0068e+002	6.5151e+001	2.1034e+002	3.7720e+002	9.1102e+001
1e5	1 st	3.0000e+002	5.3934e+002	5.5947e+002	2.0000e+002	3.9598e+002
	7 th	8.0000e+002	8.0017e+002	9.7050e+002	4.0393e+002	4.0099e+002
	13 th	8.0000e+002	8.1239e+002	1.2096e+003	4.1005e+002	4.0765e+002
	19 th	1.1626e+003	8.4319e+002	1.2626e+003	4.2683e+002	4.1659e+002
	25 th	1.2424e+003	8.7121e+002	1.2754e+003	5.0000e+002	5.0000e+002
	Mean	8.8959e+002	8.1124e+002	1.0835e+003	4.1975e+002	4.1551e+002
	Std	2.9952e+002	6.2164e+001	2.0985e+002	5.9366e+001	2.6746e+001

表 4.10 ES、PLES 與 PSGES 平均效能表現

	Traditional ES	PLES	PSGES
Function 1	9.8607e-009 (3)	8.4020e-009 (2)	0.0000e+000 (1)
Function 2	2.9000e-006 (3)	9.6502e-009 (2)	0.0000e+000 (1)
Function 3	3.5217e+005 (3)	1.1806e+005 (2)	3.1711e+000(1)
Function 4	4.1357e+003 (2)	6.0334e+003 (3)	1.3642e-014 (1)
Function 5	1.3682e+003 (2)	9.0571e+002 (3)	1.0535e+002 (1)
Function 6	7.4904e+001 (3)	3.0558e+001 (2)	1.5946e-001 (1)
Function 7	1.1826e+000 (2)	4.0943e+000 (3)	7.3913e-003 (1)
Function 8	2.0368e+001 (1)	2.0383e+001 (2)	2.0909e+001 (3)
Function 9	4.4852e+001 (3)	1.6728e+001 (2)	3.4625e+000 (1)
Function 10	1.0302e+002 (3)	2.5630e+001 (2)	1.4686e+001 (1)
Function 11	8.9174e+000 (1)	9.5263e+000 (2)	1.3546e+001 (3)
Function 12	4.4012e+003 (3)	3.2595e+003 (2)	3.6095e+002 (1)
Function 13	9.5956e+000 (3)	8.6683e+000 (2)	8.2182e-001 (1)
Function 14	3.5314e+000 (1)	4.1382e+000 (2)	5.0000e+000 (3)
Function 15	5.7165e+002 (3)	3.7977e+002 (2)	3.2644e+002 (1)
Function 16	4.3859e+002 (3)	1.4671e+002 (1)	2.0136e+002 (2)
Function 17	4.4994e+002 (3)	1.9581e+002 (1)	3.0325e+002 (2)
Function 18	1.1433e+003 (3)	1.0148e+003 (2)	7.1587e+002 (1)
Function 19	1.1250e+003 (3)	1.0019e+003 (2)	6.6960e+002 (1)
Function 20	1.1258e+003 (3)	9.9894e+002 (2)	7.0579e+002 (1)
Function 21	1.3199e+003 (3)	1.0794e+003 (2)	8.8959e+002 (1)
Function 22	9.2978e+002 (3)	8.8049e+002 (2)	8.1124e+002 (1)
Function 23	1.3495e+003 (3)	1.1141e+003 (2)	1.0835e+003 (1)
Function 24	1.1959e+003 (3)	2.8238e+002 (1)	4.1975e+002 (2)
Function 25	4.1518e+002 (1)	6.9231e+002 (3)	4.1551e+002 (2)
Rank	2.56 (64/25)	2.04 (51/25)	1.40 (35/25)

4.4.2 引導式演化策略與其他進階演算法比較

在實驗比較的第二部份，我們將所提的方法與目前進階的演算法基於這組 benchmark 下的結果做比較。此部分的比較重點，著重在觀察 PSGES 相對於其他進階演算法對於求解能力上的比較，基於這組 benchmark 25 個 test functions 下，探討可解決掉的問題個數。由表 4.3 的求解臨界值定義中，我們將問題分成三個子部份：

- Unimodal functions

25 次獨立實驗中，所得的平均值結果達到臨界值的標準，則我們稱此測試函式可解決。

- Solved multimodal functions :

針對個別測試函式，若在 25 次獨立實驗(independent run)中至少存在一次的結果達到臨界值的標準，則我們稱此測試函式可解決。

- Unsolved multimodal functions

針對個別測試函式，若在 25 次獨立實驗(independent run)中不存在任何一次的實驗結果達到臨界值的標準，則我們稱此測試函式不可被解決。

依上述所規定的判斷標準，表 4.11 為所提演算法與其他進階演算法可求解問題各數的整理。由實驗結果得知，本文所提的演算法在 unimodal 的 test function 下，可解掉 Function 1、Function 2、Function 4 三個問題；在 multimodal

的 test functions 下可解掉 Function 6、Function 7、Function 9、Function 12 與 Function 13 等 5 個問題。本文所提的方法對於這組 benchmark 共可解掉 8 個問題，在沒有特別針對參數做調整的情況下，與其他進階演算法在求解能力的比較上可達到中等以上的水準。

表 4.11 PSGES 與其他進階演算法比較

Methods	Unimodal Functions (5)	Multimodal Functions (20)			Total Solved Functions
		Basic Functions (7)	Expanded Functions (2)	Hybrid Composition Functions (11)	
PSGES	1,2,4	6,7,9,12,13	*	*	8
EDA[39]	1,2,3,4	*	*	*	4
Co-EVO[40]	1,2,3,4	7	*	*	5
DE[41]	1,2,3,4,5	6,9	*	*	7
K-PCX[42]	1,2,4	6,9,10,12	*	*	7
SPC-PNX[43]	1,2,4,5	6,7,11	*	*	7
BLX-MA[44]	1,2,4,5	9,11,12	*	*	7
DMS-L-PSO[45]	1,2,3,5	6,7,9,12	15	*	9
LR-CMA-ES[46]	1,2,3,4,5	6,7,12	*	*	8

4.5 本章總結

為了有效及公平地衡量本文所提的演算法在搜尋以及求解問題上的能力，本文一開始先說明為何要使用 CEC 2005 Special Session 所提的 25 個 test functions 作為我們衡量演算法效能的測試平台，接著說明 25 個 test functions 的特性，衡量的標準，以及我們所要比較的對象，比較的對象分成兩個部份，第一部份比較傳統 ES 與 PSGES 在搜尋效能上的差異，以觀察本文所提出的方法所改善的效能，此外我們也與同樣基於 ES 做改良的 PLES 做比較，由實驗結果可知，PSGES 在搜尋的效能上明顯優於傳統 ES 與 PLES；第二部份我們將所提的方法與其他利用此組 benchmark 作為衡量標準的進階演算法做比較，根據所定義的評判標準來比較 25 個測試函式中可求解掉的函式個數，在沒有特別針對參數做調整的情況下，本文所提的方法在效能上已有可競爭性 (comparable) 的表現。

第五章 總結與未來工作

5.1 總結

經由對 ES 與 PSO 兩者的搜尋行為上做研究與探討分析，本論文著眼於對傳統 ES 上的突變行為上做改良，經由觀察出 ES 中在突變機制上的行為與限制，也就是突變行為以亂數決定所造成的不穩定性，因此本文提出的方法為擷取 PSO 中群體間相互合作的群體智能，以此為「概念」來對於偏向隨機搜尋的 ES 進行方向上的引導，賦予 ES 中個體突變的行為參照群體智能所獲得的資訊，進而達到更具系統化的突變能力。

在第二章中，我們首先介紹了傳統 ES 的歷史背景與歷史緣由，並對於過去 ES 中相關重要的里程碑做描述。接著說明 ES 中主要的遺傳運算運作以及其演化機制，其中包括編碼形式、重組、篩選以及三種不同的突變機制作說明與探討，在我們的方法中，擷取 PSO 的搜尋概念，因此我們也對於 PSO 的基本運作機制做說明，並對於其搜尋行為上的概念作探討，最後在相關研究中，描述了近年 ES 的發展，並對於與本論文有相關的研究做概括性的描述。

在第三章為主要的方法論，我們首先對於 PSGES 的想法起源做說明，探討傳統關聯式突變的行為與限制，以及如何引入 PSO 原有的「搜尋概念」，作為突變過程中方向引導的資訊。接著描述如何基於關聯式突變與座

標軸獨立的機制上加以修改，發展了結合 PSO 方向引導與 ES 中突變強度自動調節機制的引導式突變，我們比較引導式突變與傳統突變的差異，並且說明如何結合引導式突變於傳統 ES 的原有搜尋機制，進而發展出 PSGES，最後我們描述了 PSGES 的完整架構，並對於其優缺點及相關性質做討論。

在第四章為衡量本文所提的演算法在搜尋以及求解問題上的效能，本章一開始說明為何要使用 CEC 2005 Special Session 所提的 25 個 test functions 作為我們衡量演算法的測試平台，接著說明 25 個 test functions 的特性，衡量的標準，以及我們所要比較的對象。由實驗結果得知，PSGES 不論在與傳統 ES 亦或是與其他進階演算法進行比較，在沒有特別針對參數做調整的情況下



，本文所提的引入 PSO 的引導概念後的 PSGES 在效能上已能有極具競爭性 (comparable) 的表現。

5.2 未來工作

ES 的本身參數的調整，一向是重要的研究主題，參數的設定對於演算法效能表現上的關係極為密切，本研究所提的 PSGES，在效能評估上所利用的參數設定，為利用過去在簡單的模型下，以實現性的經驗法則所推導出來，由於引進了新的機制：群體智能的想法，因此，對於其參數本身的設定也該有所調整，未來工作之一將針對不同參數設定下，對於所提演算法本身的效能影響進行分析與探討，進而推導出一套適用於 PSGES 的參數設定。

由實驗上所得的結果，觀察到本文所提的方法對於有明顯限制條件上的問題，其效能表現上不如預期的理想，其原因歸納為 ES 本身對於此類問題上的限制，未來之二將利用處理條件限制問題上的一些方法與技巧，與本研究所提方法進行整合，並對於條件限制上的問題進行探討。

ES 由發展至今，建立在一系列機率的理論模型分析，未來工作之三將針對 PSGES 的演化過程進行收斂行為上的分析與探討。

5.3 主要總結

本文主要的目的為針對傳統 ES 的突變機制上做改良，由於傳統 ES 的突變機制上，主要是以隨機亂數來決定其突變的行為，self-adaptation 雖然能夠在演化過程中適度地調整策略參數以降低其亂數擾動的性質，然而對關聯式突變中，由於 rotation angle 此部分的策略參數與問題維度呈現二次平方項次成長，造成 self-adaptation 對此部分的調整能力上有所限制，rotation angles 決定個體所能夠突變的方向，因而我們可以說傳統的突變機制，對於突變方向的決定上仍存在大量的隨機因子所造成的不穩定性。在本文中提出了一個引導式突變的機制，其主要精神為擷取 PSO 中群體間相互合作的群體智能，以此為「概念」來對於偏向隨機搜尋的 ES 進行方向上的引導，進而降低突變方向的決定上造成的不穩定性，由視覺化的實驗觀察，此引導的概念將有效地使個體在突變的過程中，達到更有具系統化的突變能力，因此能有效率地縮短實驗時間及成本，進而達到改善並穩定其搜尋能力。經由實驗結果的分析與探討，結合了引導式突變的 PSGES，於 2005 年 CEC Special Session 所提供的 benchmark 下所表現出來的結果，確實能大幅度增進傳統 ES 的搜尋效能，且與目前一些進階演算法在求解問題的能力上，在尚未對參數做特別調整上，已具有可競爭性 (comparable) 的表現。

本研究對於演化計算領域所做的主要貢獻，為將 PSO 的群體智能 (swarm

intelligence)概念與傳統 ES 中的演化機制做一個本質及概念層次上的結合，此種結合的方式保留了 ES 以 self-adaptation 來對突變行為進行調整的優勢，更可充分地發揮群體智能對於搜尋行為上所增進的效益，相較於過去僅僅在計算層次上的結合，本研究提供了另一個可加以深入研究的思維方向。



附錄一

CEC'05 25 個 test functions 之特性與定義

1. F_1 : Shifted Sphere Function

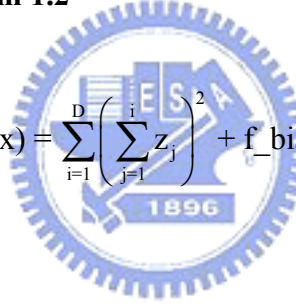
$$F_1(x) = \sum_{i=1}^D z_i^2 + f_bias_1 \quad (1)$$

$$z = x - o, x = [x_1, x_2, \dots, x_D]$$

D: dimensions. $o = [o_1, o_2, \dots, o_D]$: the shifted global optimum

2. F_2 : Shifted Schwefel's Problem 1.2

$$F_2(x) = \sum_{i=1}^D \left(\sum_{j=1}^i z_j \right)^2 + f_bias_2 \quad (2)$$



$$z = x - o, x = [x_1, x_2, \dots, x_D]$$

D: dimensions

$o = [o_1, o_2, \dots, o_D]$: the shifted global optimum

3. F_3 : Shifted Rotated High Conditioned Elliptic Function

$$F_3(x) = \sum_{i=1}^D \left(10^6 \right)^{\frac{i-1}{D-1}} z_i^2 + f_bias_3 \quad (3)$$

$$z = (x - o) * M, x = [x_1, x_2, \dots, x_D]$$

D: dimensions

$o = [o_1, o_2, \dots, o_D]$: the shifted global optimum

M: orthogonal matrix

4. F_4 : Shifted Schwefel's Problem 1.2 with Noise in Fitness

$$F_4(x) = \left(\sum_{i=1}^D \left(\sum_{j=1}^i z_j \right)^2 \right) * (1 + 0.4|N(0,1)|) + f_bias_4 \quad (4)$$

$z = x - o$, $x = [x_1, x_2, \dots, x_D]$

D = dimensions

$o = [o_1, o_2, \dots, o_D]$: the shifted global optimum

5. F_5 : Schwefel's Problem 2.6 with Global Optimum on Bounds

$$f(x) = \max \{ |x_1 + 2x_2 - 7|, |2x_1 + x_2 - 5| \}, i = 1, \dots, n, x^* = [1, 3], f(x^*) = 0$$

Extend to D dimensions:

$$F_5(x) = \max \{ |A_i x - B| \} + f_bias_5 \quad (5)$$

$i = 1, \dots, D$, $x = [x_1, x_2, \dots, x_D]$, D : dimensions

A is a $D * D$ matrix, a_{ij} are integer random number in the range $[-500, 500]$,

$\det(A) \neq 0$, A_i is the i^{th} row of A

$B_i = A_i * o$, o is a $D * 1$ vector, o_i are random number in the range $[-100, 100]$

After load the data file, set $o_i = -100$, for $i=1, 2, \dots, [D/4]$, $o_i = 100$, for $i=[3D/4], \dots, D$

Basic Multimodal Function

6. F_6 : Shifted Rosenbrock's Function

$$F_6(x) = \sum_{i=1}^{D-1} \left(100(z_i^2 - z_{i+1})^2 + (z_i - 1)^2 \right) + f_bias_6 \quad (6)$$

$z = x - o + 1$, $x = [x_1, x_2, \dots, x_D]$

D : dimensions

$o = [o_1, o_2, \dots, o_D]$: the shifted global optimum

7. F_7 : Shifted Rotated Griewank's Function without Bounds

$$F_7(x) = \sum_{i=1}^D \frac{z_i^2}{4000} - \prod_{i=1}^D \cos\left(\frac{z_i}{\sqrt{i}}\right) + 1 + f_bias_7 \quad (7)$$

$z = (x - o) * M$, $x = [x_1, x_2, \dots, x_D]$

D : dimensions

$o = [o_1, o_2, \dots, o_D]$: the shifted global optimum

M' : linear transformation matrix, condition number = 3

$M = M'(1 + 0.3|N(0,1)|)$

8. F_8 : Shifted Rotated Ackley's Function with Global Optimum on Bounds

$$F_8(x) = -20 \exp\left(-0.2 \sqrt{\frac{1}{D} \sum_{i=1}^D z_i^2}\right) - \exp\left(\frac{1}{D} \sum_{i=1}^D \cos(2\pi z_i)\right) + 20 + e + f_bias_8 \quad (8)$$

$z = (x - o) * M$,

$x = [x_1, x_2, \dots, x_D]$, D : dimensions

$o = [o_1, o_2, \dots, o_D]$: the shifted global optimum;

After load the data file, set $o_{2j-1} = -32o_{2j}$ are randomly distributed in the search range,

for $j = 1, 2, \dots, \lfloor D/2 \rfloor$

M : linear transformation matrix, condition number = 100

9. F_9 : Shifted Rastrigin's Function

$$F_9(x) = \sum_{i=1}^D (z_i^2 - 10 \cos(2\pi z_i) + 10) + f_bias_9 \quad (9)$$

$z = x - o$, $x = [x_1, x_2, \dots, x_D]$

D : dimensions

$o = [o_1, o_2, \dots, o_D]$: the shifted global optimum

10. F_{10} : Shifted Rotated Rastrigin's Function

$$F_{10}(x) = \sum_{i=1}^D (z_i^2 - 10\cos(2\pi z_i) + 10) + f_bias_{10} \quad (10)$$

$$z = (x - o) * M, x = [x_1, x_2, \dots, x_D]$$

D: dimensions

$o = [o_1, o_2, \dots, o_D]$: the shifted global optimum

M: linear transformation matrix, condition number = 2

11. F_{11} : Shifted Rotated Weierstrass Function

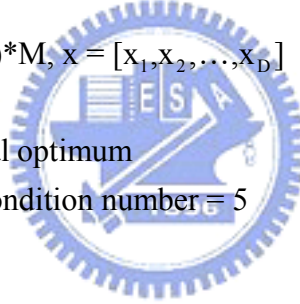
$$F_{11}(x) = \sum_{i=1}^D \left(\sum_{k=0}^{k_{\max}} \left[a^k \cos(2\pi b^k (z_i + 0.5)) \right] \right) - D \sum_{k=0}^{k_{\max}} \left[a^k \cos(2\pi b^k \cdot 0.5) \right] + f_bias_{11} \quad (11)$$

$$a = 0.5, b = 3, k_{\max} = 20, z = (x - o) * M, x = [x_1, x_2, \dots, x_D]$$

D: dimensions

$o = [o_1, o_2, \dots, o_D]$: the shifted global optimum

M: linear transformation matrix, condition number = 5



12. F_{12} : Schwefel's Problem 2.13

$$F_{12}(x) = \sum_{i=1}^D (A_i - B_i(x))^2 + f_bias_{12} \quad (12)$$

$$x = [x_1, x_2, \dots, x_D]$$

$$A_i = \sum_{j=1}^D (a_{ij} \sin \alpha_j + b_{ij} \cos \alpha_j), B_i(x) = \sum_{j=1}^D (a_{ij} \sin x_j + b_{ij} \cos x_j), \text{ for } i=1, \dots, D$$

D: dimensions

A, B are two $D \times D$ matrix, a_{ij}, b_{ij} are integer random number in the range $[-100, 100]$,

$\alpha = [\alpha_1, \alpha_2, \dots, \alpha_D]$, α_j are random numbers in the range $[-\pi, \pi]$.

Expanded Functions

Using a 2-D function $F(x, y)$ as a starting function, corresponding expanded function is:

$$EF(x_1, x_2, \dots, x_D) = F(x_1, x_2) + F(x_2, x_3) + \dots + F(x_{D-1}, x_D) + F(x_D, x_1)$$

13. F_{13} : Expanded Extended Griewank's plus Rosenbrock's Function (F8F2)

$$F8: \text{Griewank's Function: } F8(x) = \sum_{i=1}^D \frac{x_i^2}{4000} - \prod_{i=1}^D \cos\left(\frac{x_i}{\sqrt{i}}\right) + 1$$

$$F2: \text{Rosenbrock's Function: } F2(x) = \sum_{i=1}^{D-1} (100(x_i^2 - x_{i+1})^2 + (x_i - 1)^2)$$

$$F8F2(x_1, x_2, \dots, x_D) = F8(F2(x_1, x_2)) + F8(F2(x_2, x_3)) + \dots + F8(F2(x_{D-1}, x_D)) + F8(F2(x_D, x_1))$$

shift to

$$F_{13}(x) = F8(F2(z_1, z_2)) + F8(F2(z_2, z_3)) + \dots + F8(F2(z_{D-1}, z_D)) + F8(F2(z_D, z_1)) + f_bias_{13} \quad (13)$$

$$z = x - o + 1, x = [x_1, x_2, \dots, x_D]$$

D: dimensions, $o = [o_1, o_2, \dots, o_D]$: the shifted global optimum



14. F_{14} : Shifted Rotated Expanded Scaffer's F6

$$F(x, y) = 0.5 + \frac{(\sin^2(\sqrt{x^2 + y^2}) - 0.5)}{(1 + 0.001(x^2 + y^2))^2}$$

Expanded to

$$F_{14}(x) = EF(z_1, z_2, \dots, z_D) = F(z_1, z_2) + F(z_2, z_3) + \dots + F(z_{D-1}, z_D) + F(z_D, z_1) + f_bias_{14} \quad (14)$$

$$z = (x - o) * M, x = [x_1, x_2, \dots, x_D]$$

D: dimensions

$o = [o_1, o_2, \dots, o_D]$: the shifted global optimum

M: linear transformation matrix, condition number = 3

Composition functions

$F(x)$: new composition function

$f_i(x)$: i^{th} basic function used to construct the composition function

n : number of basic functions

D : dimensions

M_i : linear transformation matrix for each $f_i(x)$

o_i : new shifted optimum position for each $f_i(x)$

$$F(x) = \sum_{i=1}^n \{w_i * [f_i'((x - o_i)/\lambda_i * M_i) + \text{bias}_i]\} + f_bias$$

w_i : weight value for each $f_i(x)$, calculated as below:

$$w_i = \exp \left(- \frac{\sum_{k=1}^D (x_k - o_{ik})^2}{2D\sigma_i^2} \right),$$

$$w_i = \begin{cases} w_i & w_i = \max(w_i) \\ w_i * (1 - \max(w_i)^{10}) & w_i \neq \max(w_i) \end{cases}$$

then normalize the weight $w_i = w_i / \sum_{i=1}^n w_i$

σ_i : used to control each $f_i(x)$'s coverage range, a small σ_i give a narrow range for that $f_i(x)$

λ_i : used to stretch compress the function, $\lambda_i > 1$ means stretch, $\lambda_i < 1$ means compress

o_i define the global and local optima's position, bias_i define which optimum is global optimum.

Using o_i , bias_i , a global optimum can be placed anywhere.

If $f_i(x)$ are different functions, different functions have different properties and height,

in order to get a better mixture, estimate a biggest function value $f_{\max_i}$ for 10 function $f_i(x)$,

then normalize each basic functions to similar heights as below:

$f_i'(x) = C * f_i(x) / |f_{\max_i}|$, C is a predefined constant.

$|f_{\max_i}|$ is estimated using $|f_{\max_i}| = f_i((x'/\lambda_i) * M_i)$, $x' = [5, 5, \dots, 5]$.

In the following composition functions,

Number of basic functions $n=10$.

D: dimensions

σ : $n * D$ matrix, defines $f_i(x)$'s global optimal positions

bias = [0, 100, 200, 300, 400, 500, 600, 700, 800, 900].

Hence, the first function $f_1(x)$ always the function with the global optimum

$C=2000$

15. F_{15} : Hybrid Composition Function

$f_{1-2}(x)$: Rastrigin's Function

$$f_i(x) = \sum_{i=1}^D (x_i^2 - 10\cos(2\pi x_i) + 10)$$

$f_{3-4}(x)$: Weierstrass Function

$$f_i(x) = \sum_{i=1}^D \left(\sum_{k=0}^{k_{\max}} \left[a^k \cos(2\pi b^k (x_i + 0.5)) \right] \right) - D \sum_{k=0}^{k_{\max}} \left[a^k \cos(2\pi b^k \cdot 0.5) \right],$$

$a=0.5$, $b=3$, $k_{\max}=20$

$f_{5-6}(x)$: Griewank's Function

$$f_i(x) = \sum_{i=1}^D \frac{x_i^2}{4000} - \prod_{i=1}^D \cos\left(\frac{x_i}{\sqrt{i}}\right) + 1$$

$f_{7-8}(x)$: Ackley's function

$$f_i(x) = -20 \exp\left(-0.2 \sqrt{\frac{1}{D} \sum_{i=1}^D x_i^2}\right) - \exp\left(\frac{1}{D} \sum_{i=1}^D \cos(2\pi x_i)\right) + 20 + e$$

$f_{9-10}(x)$: Sphere Function

$$f_i(x) = \sum_{i=1}^D x_i^2$$

$\sigma_i=1$ for $i = 1, 2, \dots, D$

$\lambda = [1, 1, 10, 10, 5/60, 5/60, 5/32, 5/32, 5/100, 5/100]$

M_i are all identity matrices

16. F_{16} : Rotated Hybrid Composition Function

Except M_i are different linear transformation matrixes with condition number of 2, all other setting are the same as F_{15} .

17. F_{17} : Rotated Hybrid Composition Function with Noise in Fitness

Let $(F_{16} - f_{\text{bias}_{16}})$ be $G(x)$, then

$$F_{17}(x) = G(x) * (1 + 0.2|N(0, 1)|) + f_{\text{bias}_{17}} \quad (15)$$

All setting are the same as F_{16} .

18. F_{18} : Rotated Hybrid Composition Function

$f_{1-2}(x)$: Ackley's Function

$$f_i(x) = -20 \exp \left(-0.2 \sqrt{\frac{1}{D} \sum_{i=1}^D x_i^2} \right) - \exp \left(\frac{1}{D} \sum_{i=1}^D \cos(2\pi x_i) \right) + 20 + e$$

$f_{3-4}(x)$: Rastrigin's Function

$$f_i(x) = \sum_{i=1}^D (x_i^2 - 10 \cos(2\pi x_i) + 10)$$

$f_{5-6}(x)$: Sphere Function

$$f_i(x) = \sum_{i=1}^D x_i^2$$

$f_{7-8}(x)$: Weierstrass Function

$$f_i(x) = \sum_{i=1}^D \left(\sum_{k=0}^{k_{\max}} \left[a^k \cos(2\pi b^k (x_i + 0.5)) \right] \right) - D \sum_{k=0}^{k_{\max}} \left[a^k \cos(2\pi b^k \cdot 0.5) \right],$$

$$a=0.5, b=3, k_{\max}=20$$

$f_{9-10}(x)$: Griewank's Function

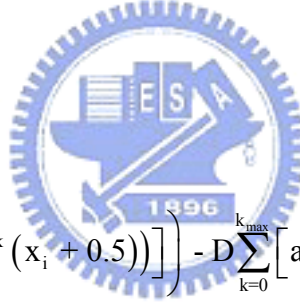
$$f_i(x) = \sum_{i=1}^D \frac{x_i^2}{4000} - \prod_{i=1}^D \cos \left(\frac{x_i}{\sqrt{i}} \right) + 1$$

$$\sigma = [1, 2, 1.5, 1.5, 1, 1, 1.5, 1.5, 2, 2];$$

$$\lambda = [2 * 5/32; 5/32; 2 * 1; 1; 2 * 5/100; 5/100; 2 * 10; 10; 2 * 5/60; 5/60]$$

M_i are all rotation matrices. condition number are [2 3 2 3 2 3 20 30 200 300]

$$o_{10}[0,0,\dots,0]$$



19. F_{19} : Rotated Hybrid Composition Function with a narrow basin for the global optimum

All settings are the same as F_{18} except

$$\sigma = [0.1, 2, 1.5, 1.5, 1, 1, 1.5, 1.5, 2, 2];$$

$$\lambda = [0.1 * 5/32; 5/32; 2 * 1; 1; 2 * 5/100; 5/100; 2 * 10; 10; 2 * 5/60; 5/60]$$

20. F_{20} : Rotated Hybrid Composition Function with the Global Optimum on the Bounds

All settings are the same as F_{18} except after load the data file, set $\sigma_{1(2j)} = 5$, for $j=1, 2, \dots, \lfloor D/2 \rfloor$

21. F_{21} : Rotated Hybrid Composition Function

$f_{1-2}(x)$: Rotated Expanded Scaffer's F6 Function

$$F(x, y) = 0.5 + \frac{\left(\sin^2 \left(\sqrt{x^2 + y^2} \right) - 0.5 \right)}{\left(1 + 0.001(x^2 + y^2) \right)^2}$$

$$f_i(x) = F(x_1, x_2) + F(x_2, x_3) + \dots + F(x_{D-1}, x_D) + F(x_D, x_1)$$

$f_{3-4}(x)$: Rastrigin's Function

$$f_i(x) = \sum_{i=1}^D \left(x_i^2 - 10 \cos(2\pi x_i) + 10 \right)$$

$f_{5-6}(x)$: F8F2 Function

$$F8(x) = \sum_{i=1}^D \frac{x_i^2}{4000} - \prod_{i=1}^D \cos\left(\frac{x_i}{\sqrt{i}}\right) + 1$$

$$F2(x) = \sum_{i=1}^{D-1} \left(100(x_i^2 - x_{i+1}^2) + (x_i - 1)^2 \right)$$

$$f_i(x) = F8(F2(x_1, x_2)) + F8(F2(x_2, x_3)) + \dots + F8(F2(x_{D-1}, x_D)) + F8(F2(x_D, x_1))$$

$f_{7-8}(x)$: Weierstrass Function

$$f_i(x) = \sum_{i=1}^D \left(\sum_{k=0}^{k_{\max}} \left[a^k \cos(2\pi b^k (x_i + 0.5)) \right] \right) - D \sum_{k=0}^{k_{\max}} \left[a^k \cos(2\pi b^k \cdot 0.5) \right],$$

$$a=0.5, b=3, k_{\max}=20$$

$f_{9-10}(x)$: Griewank's Function

$$f_i(x) = \sum_{i=1}^D \frac{x_i^2}{4000} - \prod_{i=1}^D \cos\left(\frac{x_i}{\sqrt{i}}\right) + 1$$

$$\sigma = [1, 1, 1, 1, 1, 2, 2, 2, 2, 2],$$

$$\lambda = [5 * 5 / 100; 5 / 100; 5 * 1; 1; 5 * 1; 1; 5 * 10; 10; 5 * 5 / 200; 5 / 200];$$

M_i are all orthogonal matrix

22. F_{22} : Rotated Hybrid Composition Function with High Condition Number Matrix

All settings are the same as F_{21} except M_i 's condition number are
[10 20 50 100 200 1000 2000 3000 4000 5000]

23. F_{23} : Non-Continuous Rotated Hybrid Composition Function

All settings are the same as F_{21}

$$\text{Except } x_j = \begin{cases} x_j & |x_j - o_{ij}| < 1/2 \\ \text{round}(2x_j)/2 & |x_j - o_{ij}| \geq 1/2 \end{cases} \text{ for } j = 1, 2, \dots, D$$

$$\text{round}(x) = \begin{cases} a - 1 & \text{if } x \leq 0 \text{ and } b \geq 0.5 \\ a & \text{if } b < 0.5 \\ a + 1 & \text{if } x > 0 \text{ and } b \geq 0.5 \end{cases}$$

where a is x 's integral part and b is x 's decimal part

All "round" operators use the same schedule.

24. F_{24} : Rotated Hybrid Composition Function

$f_1(x)$: Weierstrass Function

$$f_i(x) = \sum_{i=1}^D \left(\sum_{k=0}^{k_{\max}} \left[a^k \cos(2\pi b^k (x_i + 0.5)) \right] \right) - D \sum_{k=0}^{k_{\max}} \left[a^k \cos(2\pi b^k \cdot 0.5) \right],$$

$$a=0.5, b=3, k_{\max}=20$$

$f_2(x)$: Rotated Expanded Scaffer's F6 Function

$$F(x, y) = 0.5 + \frac{\left(\sin^2 \left(\sqrt{x^2 + y^2} \right) - 0.5 \right)}{\left(1 + 0.001(x^2 + y^2) \right)^2}$$

$$f_i(x) = F(x_1, x_2) + F(x_2, x_3) + \dots + F(x_{D-1}, x_D) + F(x_D, x_1)$$

$f_3(x)$: F8F2 Function

$$F8(x) = \sum_{i=1}^D \frac{x_i^2}{4000} - \prod_{i=1}^D \cos\left(\frac{x_i}{\sqrt{i}}\right) + 1$$

$$F2(x) = \sum_{i=1}^{D-1} \left(100(x_i^2 - x_{i+1})^2 + (x_i - 1)^2 \right)$$

$$f_i(x) = F8(F2(x_1, x_2)) + F8(F2(x_2, x_3)) + \dots + F8(F2(x_{D-1}, x_D)) + F8(F2(x_D, x_1))$$

$f_4(x)$: Ackley's Function

$$f_i(x) = -20 \exp\left(-0.2 \sqrt{\frac{1}{D} \sum_{i=1}^D x_i^2}\right) - \exp\left(\frac{1}{D} \sum_{i=1}^D \cos(2\pi x_i)\right) + 20 + e$$

$f_5(x)$: Rastrigin's Function

$$f_i(x) = \sum (x_i^2 - 10 \cos(2\pi x_i) + 10)$$

$f_6(x)$: Griewank's Function

$$f_i(x) = \sum_{i=1}^D \frac{x_i^2}{4000} - \prod_{i=1}^D \cos\left(\frac{x_i}{\sqrt{i}}\right) + 1$$

$f_7(x)$: Non-Continuous Expanded Scaffer's F6 Function

$$F(x, y) = 0.5 + \frac{\left(\sin^2(\sqrt{x^2 + y^2}) - 0.5\right)}{\left(1 + 0.001(x^2 + y^2)\right)^2}$$

$$f(x) = F(y_1, y_2) + F(y_2, y_3) + \dots + F(y_{D-1}, y_D) + F(y_D, y_1)$$

$$y_j = \begin{cases} x_j & |x_j| < 1/2 \\ \text{round}(2x_j)/2 & |x_j| \geq 1/2 \end{cases} \quad \text{for } j=1, 2, \dots, D$$

$f_8(x)$: Non-Continuous Rastrigin's Function

$$f(x) = \sum_{i=1}^D (y_i^2 - 10 \cos(2\pi y_i) + 10)$$

$$y_j = \begin{cases} x_j & |x_j| < 1/2 \\ \text{round}(2x_j)/2 & |x_j| \geq 1/2 \end{cases} \quad \text{for } j=1, 2, \dots, D$$

$f_9(x)$: High Conditioned Elliptic Function

$$f(x) = \sum_{i=1}^D \left(10^6\right)^{\frac{i-1}{D-1}} x_i^2$$

$f_{10}(x)$: Sphere Function with Noise in Fitness

$$f_i(x) = \left(\sum_{i=1}^D x_i^2 \right) (1 + 0.1 |N(0, 1)|)$$

$\sigma_i=2$, for $i=1,2,\dots,D$

$\lambda=[10; 5/20; 1; 5/32; 1; 5/100; 5/50; 1; 5/100; 5/100]$

M_i are all rotation matrices, condition numbers are $[100\ 50\ 30\ 10\ 5\ 5\ 4\ 3\ 2\ 2]$;

25. *F*₂₅: Rotated Hybrid Composition Function without Bounds

All settings are the same as *F*₂₄ except no exact search range set for this test function.



附錄二

CEC'05 25 個 test functions 整體最佳解的個 體目標變數與其適應值

Function	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8	X_9	X_{10}	$f(x)$
1	-39.312	58.900	-46.322	-74.651	-16.8	-80.544	-10.594	24.969	9.838	9.1119	-450
2	35.627	-82.912	-10.642	-83.582	83.155	47.048	-89.436	-27.422	76.145	-39.06	-450
3	-32.201	64.978	-38.3	-23.258	-54.009	86.629	-6.3009	-49.364	5.3499	52.242	-450
4	35.627	-82.912	-10.642	-83.582	83.155	47.048	-89.436	-27.422	76.145	-39.06	-450
5	-100	-100	-100	8.3897	7.7182	-8.3147	100	100	100	100	-310
6	81.023	-48.395	19.232	-2.5231	70.434	47.177	-7.8358	-86.669	57.853	-9.9533	390
7	-276.27	-11.911	-578.79	-287.65	-84.386	-228.68	-458.15	-202.21	-105.86	-96.49	-180
8	-32	14.977	-32	9.5566	-32	-17.19	-32	0.8511	-32	10.793	-140
9	1.9005	-1.5644	-0.9788	-2.2536	2.499	-3.2853	0.9759	-3.6661	0.0985	-3.2465	-330
10	1.9005	-1.5644	-0.9788	-2.2536	2.499	-3.2853	0.9759	-3.6661	0.0985	-3.2465	-330
11	-0.1367	0.1186	-0.0968	0.0237	-0.2933	-0.0478	0.3518	0.3579	-0.0586	-0.0375	90
12	-2.028	-1.5589	0.7774	-2.0752	-0.1601	1.0811	1.408	-1.6129	2.419	2.217	-460
13	0.2471	-0.8497	0.5629	0.2673	0.6874	0.1395	-0.2764	0.4172	-0.5007	-0.5204	-130
14	-73.603	-23.55	-21.774	18.134	32.724	34.413	80.946	48.147	19.151	10.103	-300
15	3.3253	-1.2835	1.8984	-0.4095	0.0881	2.758	0.9776	-1.809	-2.4957	2.7367	120
16	3.3253	-1.2835	1.8984	-0.4095	0.0881	2.758	0.9776	-1.809	-2.4957	2.7367	120
17	3.3253	-1.2835	1.8984	-0.4095	0.0881	2.758	0.9776	-1.809	-2.4957	2.7367	120
18	1.5953	2.644	1.8047	0.9389	-3.0486	-1.1571	3.5582	2.4246	-0.3767	4.4637	10
19	1.5953	2.644	1.8047	0.9389	-3.0486	-1.1571	3.5582	2.4246	-0.3767	4.4637	10
20	1.5953	5	1.8047	5	-3.0486	5	3.5582	5	-0.3767	5	10
21	1.2141	-0.01	1.8864	-4.1124	2.0627	1.1535	4.0653	-1.0213	1.1986	-4.1795	360
22	1.2141	-0.01	1.8864	-4.1124	2.0627	1.1535	4.0653	-1.0213	1.1986	-4.1795	360
23	1.2141	-0.01	1.8864	-4.1124	2.0627	1.1535	4.0653	-1.0213	1.1986	-4.1795	360
24	-2.933	-3.034	-4.3583	-4.2851	-3.505	-1.4618	-3.7277	-2.1445	-3.1317	-2.5637	260
25	-2.933	-3.034	-4.3583	-4.2851	-3.505	-1.4618	-3.7277	-2.1445	-3.1317	-2.5637	260

附錄三

PSGES 所達到的最佳解之個體目標變數與其適應值

Function	X ₁	X ₂	X ₃	X ₄	X ₅	X ₆	X ₇	X ₈	X ₉	X ₁₀	f(x)
1	-39.312	58.9	-46.322	-74.651	-16.8	-80.544	-10.594	24.969	89.838	9.1119	-450
2	35.627	-82.912	-10.642	-83.581	83.155	47.048	-89.436	-27.422	76.145	-39.059	-450
3	-32.202	64.972	-38.305	-23.254	-54.008	86.624	-6.3001	-49.365	5.3484	52.244	-449.9997
4	35.627	-82.912	-10.642	-83.581	83.155	47.048	-89.436	-27.422	76.145	-39.059	-450
5	-98.621	-100.47	-100.96	9.1238	7.5893	-8.2788	99.773	99.92	100.15	100.26	-306.1513
6	81.023	-48.395	19.232	-2.5231	70.434	47.177	-7.8358	-86.669	57.853	-9.9533	390
7	-276.27	-11.911	-578.79	-287.65	-84.386	-228.68	-458.15	-202.21	-105.86	-96.49	-180
8	148.22	64.426	131.93	-15.17	-134.42	169.83	-33.84	136.36	54.566	-164.4	-119.2047
9	1.9005	-1.5644	-0.9788	-2.2536	2.499	-3.2853	0.9759	-3.6661	0.0985	-3.2465	-330
10	1.6761	-1.896	-0.5549	-2.7395	3.1458	-2.5447	1.0481	-3.7332	-0.4833	-3.7566	-325.0252
11	-1.2241	-5.4806	0.0062	9.1931	-0.3377	1.1227	-3.0906	4.6003	1.9495	0.6940	100.4350
12	28.729	18.728	13.848	-27.814	5.9342	45.934	94.103	-20.666	51.941	83.913	-460
13	0.2471	-0.8496	0.5629	0.2673	0.6873	0.1394	-0.2764	0.4172	-0.5007	-0.5203	-130
14	-30.866	-24.227	-55.578	22.898	13.774	23.15	62.072	68.05	51.405	-12.73	-296.1653
15	4.3258	-1.2844	1.8966	-0.4018	0.0865	3.7498	1.9703	-2.8003	-2.5032	2.734	198.5147
16	-2.8095	4.2572	-0.3487	-3.8883	-2.6367	-4.5978	2.1842	2.4808	-1.1638	-2.4208	239.7859
17	4.8851	0.3257	1.131	0.6134	-0.0592	3.0273	2.4899	-1.9942	-4.041	3.4901	303.5610
18	2.8727	-0.8312	-4.1507	-4.3428	-4.2097	-3.95	-3.7506	-3.1893	3.4647	0.4615	310
19	2.8727	-0.8312	-4.1507	-4.3428	-4.2097	-3.95	-3.7506	-3.1893	3.4647	0.4615	310
20	2.8727	-0.8312	-4.1507	-4.3428	-4.2097	-3.95	-3.7506	-3.1893	3.4647	0.4615	310
21	-2.6639	-2.1802	-3.1792	2.0201	2.4322	1.5009	-0.7393	-0.2911	-1.4312	2.7447	660
22	-8.327	14.859	42.057	3.0181	24.143	-7.1897	-17.964	-13.582	15.755	61.316	889.3623
23	-4.1802	2.0668	1.7308	3.0332	-1.4411	-1.0616	3.4114	-3.3833	3.0469	3.4588	919.4683
24	0.7973	-2.6471	4.2808	-3.1835	1.7367	-2.3262	-0.8345	-2.9156	3.6308	3.8112	460
25	-5.9106	-8.3375	15.801	-8.6864	11.07	-6.0961	-2.3272	-1.99	16.058	15.939	641.5032

參考文獻

- [1] C. Darwin, *On the origin of species*. London: John Murray, 1859.
- [2] J. H. Holland, *Adaptation in Natural and Artificial System*. Ann Arbor, Michigan: The University of Michigan Press, 1975.
- [3] D. B. Fogel, *Artificial Intelligence Through Simulated Evolution*. New York: Wiley Publishing, 1966.
- [4] I. Rechenberg, *Evolutionsstrategie: Optimierung Technischer Systeme nach Prinzipien der Biologischen Evolution*. Suttgart: Frommann-Holzboog, 1973.
- [5] T. Bäck, *Evolutionary Algorithms in Theory and Practice: Evolution Strategies, Evolutionary Programming, Genetic Algorithms*. New York: Oxford University Press, 1996.
- [6] T. Bäck, U. Hammel, and H.-P. Schwefel, "Evolutionary computation. Comments on the history and current state," *IEEE Transactions on Evolutionary Computation*, vol. 1, pp. 3-17, 1997.
- [7] I. Rechenberg, "Case studies in evolutionary experimentation and computation," *Computer Methods in Applied Mechanics and Engineering*, vol. 186, pp. 125-140, 2000.
- [8] H.-P. Schwefel, *Numerical Optimization of Computer Models*. Chichester: Wiley, 1981.
- [9] G. Rudolph, "On Correlated Mutations in Evolution Strategies," presented at Parallel Problem Solving from Nature, 1992.
- [10] J. Kennedy and R. Eberhart, "Particle swarm optimization," *Proceedings of IEEE International Conference on Neural Networks*, pp. 1942-1948, 1995.
- [11] J. Kennedy and R. C. Eberhart, "A new optimizer using particle swarm theory," *Proceeding of the Sixth Int. Symposium on Micromachine and Human Science*, pp. 39-43, 1995.
- [12] K. E. Parsopoulos and M. N. Vrahatis, "Recent approaches to global optimization problems through Particle Swarm Optimization," *Natural Computing*, vol. 1, pp. 235-306, 2002.
- [13] D. V. Arnold and H.-G. Beyer, "Performance analysis of evolution strategies with multi-recombination in high-dimensional R^N -search spaces disturbed by noise," *Theoretical Computer Science*, vol. 289, pp. 629-647, 2002.
- [14] D. V. Arnold and H.-G. Beyer, "Local Performance of the (1+1)-ES in a Noisy Environment," *IEEE Trans. Evolutionary Computation*, vol. 6, pp. 30-41, 2002.
- [15] H.-P. Schwefel, *Evolution and Optimum Seeking*. New York: John Wiley and Sons, 1995.

- [16] H.-G. Beyer, *The Theory of Evolution Strategies*. Berlin: Springer, 2001.
- [17] H.-G. Beyer, "On the performance of $(1,\lambda)$ -evolution strategies for the ridge function class," *IEEE Transaction on Evolutionary Computation*, vol. 5, pp. 218-235, 2001.
- [18] H.-G. Beyer, *The Theory of Evolution Strategies*. Heidelberg: Springer Verlag, 2001.
- [19] H.-G. Beyer, "Toward a theory of evolution strategies: On the benefit of sex - the $(\mu/\mu,\lambda)$ -theory," *Evolutionary Computation*, vol. 3, pp. 81-111, 1995.
- [20] A. Ostermeier, A. Gawelczyk, and N. Hansen, "A derandomized approach to self-adaptation of evolution strategies," *Evolutionary Computation*, vol. 2, pp. 369-380, 1994.
- [21] N. Hansen and A. Ostermeier, "Adapting arbitrary normal mutation distributions in evolution strategies: The covariance matrix adaptation," *Proceedings of the 1996 IEEE International Conference on Evolutionary Computation*, pp. 312-317, 1996.
- [22] N. Hansen and A. Ostermeier, "Completely Derandomized Self-Adaptation in Evolution Strategies," *Evolutionary Computation*, vol. 9, pp. 159-195, 2001.
- [23] L. Hidebrand, B. Reusch, and M. Fathi, "Directed mutation-a new self-adaptation for evolution strategies," *IEEE Congress on Evolutionary Computation*, vol. 2, pp. 1550-1557, 1999.
- [24] V. Miranda and N. Fonseca, "New Evolutionary Particle Swarm Algorithm (EPSO) Applied To Voltage/Var Control," in *14th Power Systems Computation Conference*. Sevilla, 2002.
- [25] V. Miranda and N. Fonseca, "EPSO - best-of-two-worlds meta-heuristic applied to power system problems," *IEEE Congress on Evolutionary Computation*, vol. 2, pp. 1080-1085, 2002.
- [26] M. Herdy, "Reproductive isolation as strategy parameters in hierarchically organized evolution strategies," *Parallel Problem Solving from Nature 2*, pp. 207-217, 1992.
- [27] N. Hansen and A. Ostermeier, "Adapting arbitrary normal mutation distributions in evolution strategies: The covariance matrix adaptation," *Proceedings of the IEEE international Conference on Evolutionary Computation*, pp. 312-317, 1996.
- [28] H.-G. Beyer, "Toward a theory of evolution strategies: Self-adaptation," *Evolutionary Computation*, vol. 3, pp. 311-347, 1996.
- [29] H.-G. Beyer, "Toward a Theory of Evolution Strategies: Some Asymptotical Results from the $(1+\lambda)$ -Theory," *Evolutionary Computation*, vol. 1, pp. 165-188, 1993.
- [30] I. Rechenberg, *Evolutions-strategie '94*. Stuttgart: frommann-holzboog, 1994.
- [31] H.-G. Beyer and H.-P. Schwefel, "Evolution strategies: A comprehensive introduction," *Natural Computing*, vol. 1, pp. 3-52, 2002.
- [32] T. Bäck, D. B. Fogel, and Z. Michalewicz, *Handbook of Evolutionary Computation*. New York: Oxford University Press, 1997.
- [33] T. Bäck, U. Hammel, and H.-P. Schwefel, "Evolutionary computation: Comments on the history and current state," *IEEE Transactions on Evolutionary Computation*, vol.

- 1, pp. 3-17, 1997.
- [34] H.-P. Schwefel, "Numerische Optimierung von Computer-Modellen mittels der Evolutionsstrategie," *Interdisciplinary Systems Research*, vol. 26, 1977.
 - [35] Y. Shi and R. Eberhart, "A modified particle swarm optimization," *IEEE International Conference on Evolutionary Computation*, pp. 69-73, 1998.
 - [36] Y. Shi and R. C. Eberhart, "Parameter selection in particle swarm optimization," *Evolutionary Programming of Lecture Notes in Computer Science*, vol. 1447, pp. 591-600, 1998.
 - [37] P. N. Suganthan, N. Hansen, J. J. Liang, K. Deb, Y.-P. Chen, A. Auger, and S. Tiwari, "Problem Definitions and Evaluation Criteria for the CEC 2005 Special Session on Real-Parameter Optimization," Nanyang Technological University, Singapore May 2005 AND KanGAL Report #2005005, IIT Kanpure, India 2005.
 - [38] L. Costa, "A Parameter-less Evolution Strategy for Global Optimization," Universidade do Minho, Escola de Engenharia, Portugal 2005.
 - [39] B. Yuan and M. Gallagher, *IEEE Congress on Evolutionary Computation*, vol. 2, pp. 1792-1799, 2005.
 - [40] P. Posik, "Real Parameter Optimization Using Mutation Step Co-evolution," *IEEE Congress on Evolutionary Computation*, vol. 1, pp. 872-879, 2005.
 - [41] J. Ronkkonen, S. Kukkonen, and K. V. Price, "Real-Parameter Optimization with Differential Evolution," *IEEE Congress on Evolutionary Computation*, vol. 1, pp. 506-513, 2005.
 - [42] A. Sinha, S. Tiwari, and K. Deb, "A Population-Based, Steady-State Procedure for Real-Parameter Optimization," *IEEE Congress on Evolutionary Computation*, vol. 1, pp. 514-521, 2005.
 - [43] P. J. Ballester, J. Stephenson, J. N. Carter, and K. Gallagher, "Real-Parameter Optimization Performance Study on the CEC-2005 benchmark with SPC-PNX," *IEEE Congress on Evolutionary Computation*, vol. 1, pp. 498-505, 2005.
 - [44] D. Molina, F. Herrera, and M. Lozano, "Adaptive Local Search Parameters for Real-Coded Memetic Algorithm," *IEEE Congress on Evolutionary Computation*, vol. 1, pp. 888-895, 2005.
 - [45] J. J. Liang and P. N. Suganthan, "Dynamic Multi-Swarm Particle Swarm Optimizer with Local Search," *IEEE Congress on Evolutionary Computation*, vol. 1, pp. 522-528, 2005.
 - [46] A. Auger and N. Hansen, "Performance Evaluation of an Advanced Local Search Evolutionary Algorithm," *IEEE Congress on Evolutionary Computation*, vol. 2, pp. 1777-1784, 2005.