RAPIDLY-EXPLORING RANDOM TREES (RRT)

David Esparza Alba

Objetivo

- Añadir la estructura de un RRT a un algoritmo de estimación de distribución (EDA) para resolver problemas de optimización.

¿Qué es un RRT?

- Es una estructura de datos diseñada para realizar búsquedas eficientes en espacios multidimensionales no convexos.
- Los RRTs pueden ser considerados como una técnica para generar trayectorias para sistemas no lineales con restricciones.

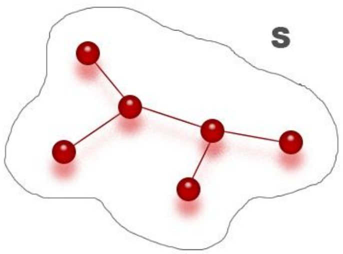
Propiedades de los RRT

- Los RRT iterativamente adapta el campo de búsqueda a una región de Voronoi de cada nodo durante el proceso de búsqueda.
- La exploración es determinada por el diagrama de Voronoi de los nodos pertenecientes al árbol. La probabilidad de que un nodo sea elegido para la expansión del árbol es proporcional al volumen de su región de Voronoi. Es por eso que un RRT tiende a crecer rápidamente a regiones inexploradas del espacio de búsqueda.

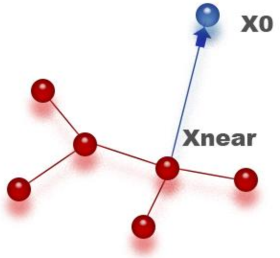
Algoritmo

1. Tomar un nodo inicial aleatorio y añadirlo a S.
2. Crear un punto aleatorio cualquiera X_0.
3. Buscar el nodo más cercano (X_{near}) a X_0 dentro de S.
4. Trazar un vector unitario de X_{near} a X_0.
   - $$A = \frac{X_0 - X_{near}}{\|X_0 - X_{near}\|}$$
5. Multiplicar este vector unitario **A** por un valor ε y lo añadimos a S.
6. Si deseamos generar mas nodos, regresamos (2).

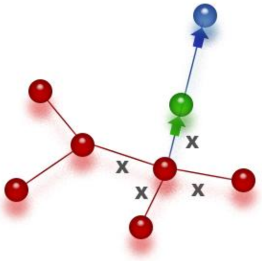
Construcción de un RRT



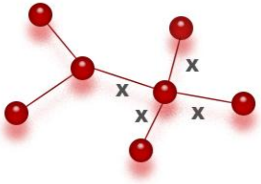
Paso 1: Identificar el conjunto S.



Paso 2: Tomar un punto X0 proveniente de una distribución y buscar el nodo (Xnear) en S más cercano a esté.

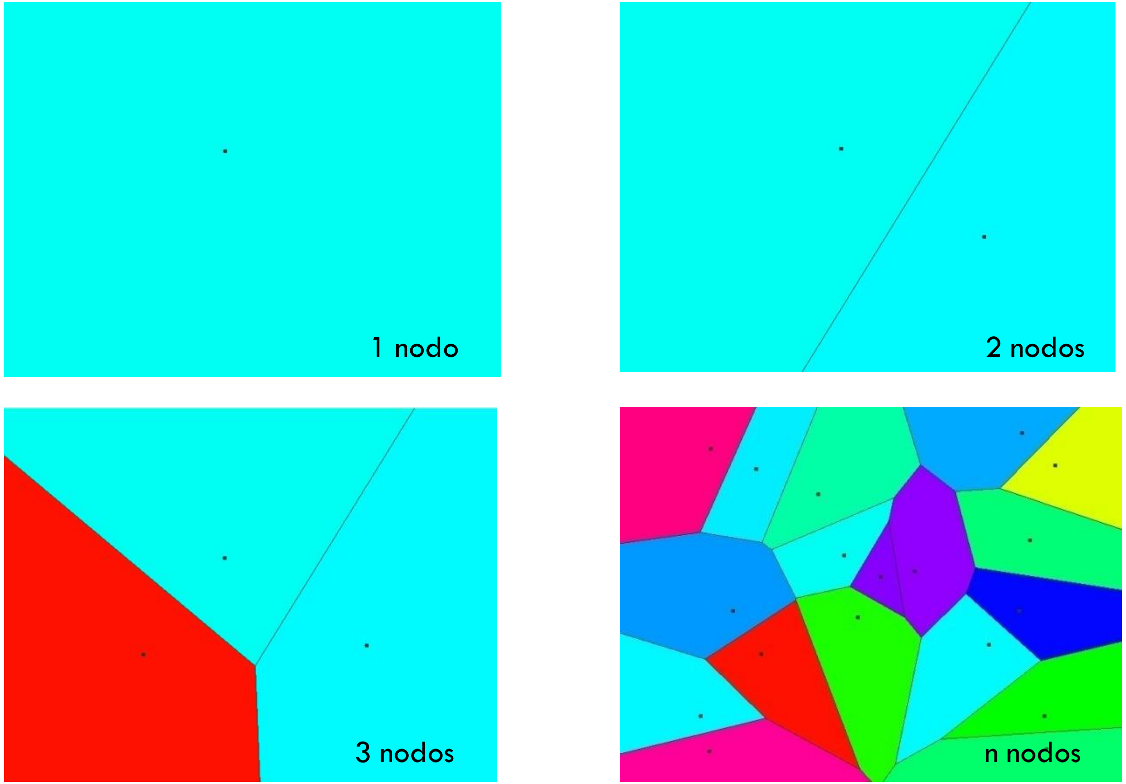


Paso 3: Trazar un vector de Magnitud “X” de Xnear a X0.



Paso 4: Agregar el nuevo nodo al conjunto S.

Construcción de un RRT



Construcción de un RRT

EDA

- Un Algoritmo de Estimación de Distribución (EDA) es un proceso iterativo. En cada iteración el EDA estima la distribución de la población por medio de una muestra, y de esta manera poder generar los nuevos individuos de la población.

EDA

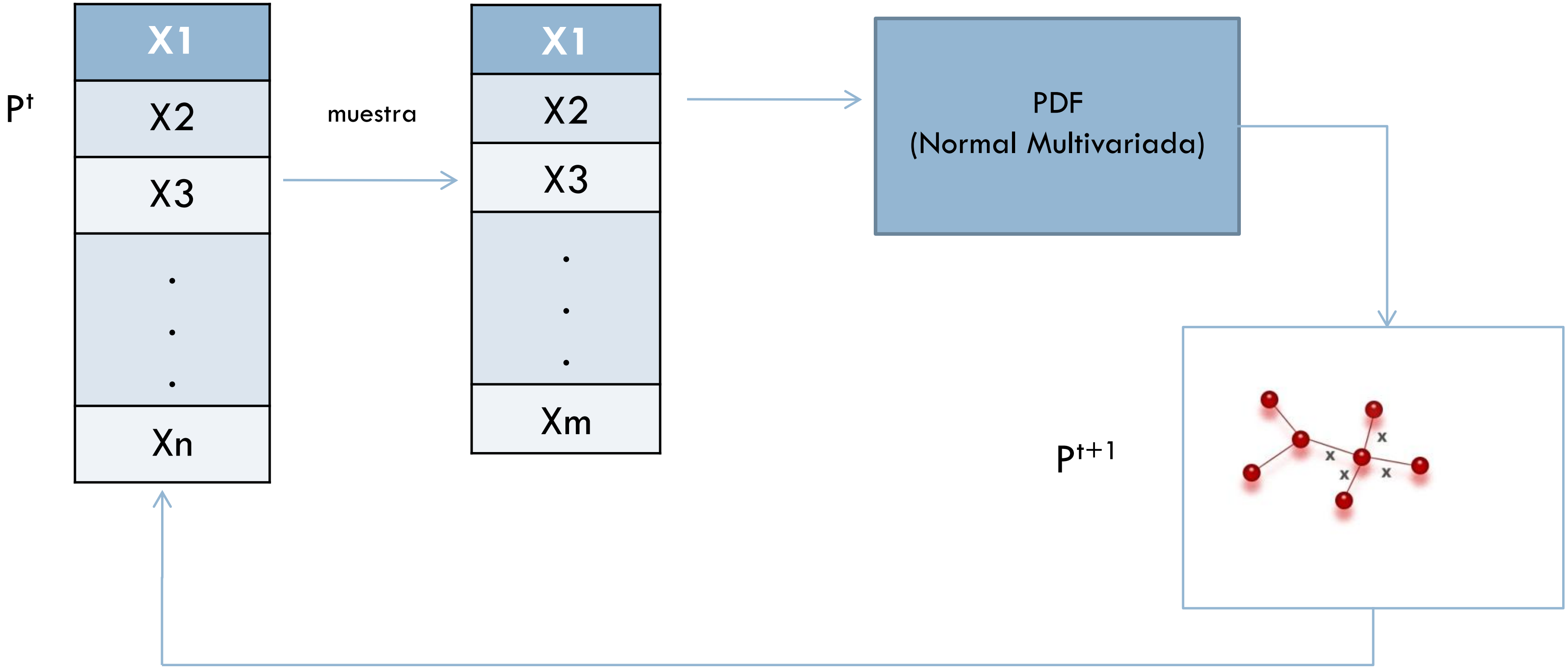
1. Generar aleatoriamente una población inicial P(0),
2. t = 0.
3. Tomar una muestra M de P.
4. Por medio de un modelo (PDF), estimar la distribución de M.
5. Generar P(t+1) con la distribución obtenida en el paso anterior.
6. t = t+1, Regresar al paso 3.

Función de Densidad de Probabilidad (PDF)

- Vamos a usar como modelo para estimar la distribución de la población una distribución normal multivariada. Para esto tenemos el siguiente algoritmo:

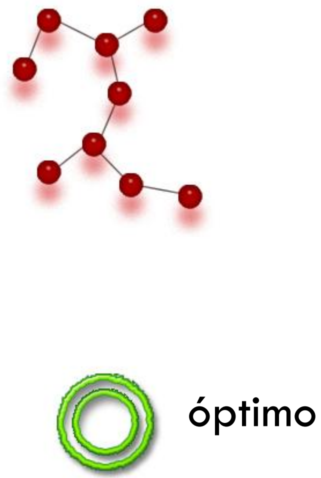
1. Tomar una muestra de buenos individuos de la población M.
2. Obtenemos el vector de medias $\boldsymbol{\mu}$ y la matriz de covarianzas $\boldsymbol{\Sigma}$.
3. Factorizar Σ en la forma $LL^{\dagger}$ (Cholesky).
4. Generamos una matriz **H** con valores aleatorios provenientes de una distribución normal con $\mu = 0$ y $\sigma^2 = 1$.
5. La nueva población la tomamos como $P(t+1) = \boldsymbol{\mu} + \mathbf{HL}^{\dagger}$

RRT + EDA



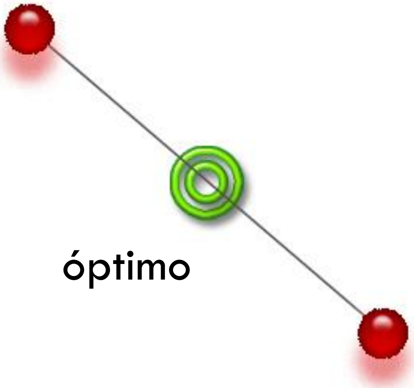
Problemas en los RRT

1. Tamaño de paso demasiado chico



La población avanza demasiado lento que no alcanza a llegar al óptimo

2. Tamaño de paso demasiado grande



El tamaño de paso puede ser demasiado grande y saltarse el óptimo y nunca encontrarlo.

Solución

- Para solucionar el problema del tamaño de paso, vamos a usar una regla simple.
- Si cada k generaciones debemos preguntar los siguiente:
- Si existe una mejora en la solución:
  - Incrementamos el tamaño de paso, multiplicándolo por un factor α ($\alpha > 1$).

  Si no existe una mejora en la solución:
  - Reducir el tamaño de paso, dividiéndolo por un factor α.

Resultados

- A continuación presentamos los resultados obtenidos corriendo el algoritmo 20 veces, los parámetros utilizados fueron los siguientes:
  - Tamaño de población: 300
  - Tamaño de muestra: 150
  - Número de evaluaciones de función: 450,000
  - Tamaña de paso: 1.1

Resultados

	Intervalo	Mejor	Peor	Media	Mediana	Desviación Estándar
DIM 5						
Sphere	$-600 < x_i < 600$	3.12633E-06	1.26042E-05	8.14442E-06	8.19858E-06	2.83896E-06
Griewank	$-600 < x_i < 600$	1.61072E-06	9.86357E-03	1.85181E-03	3.61330E-06	3.82460E-03
Ackley	$-10 < x_i < 10$	2.41631E-03	1.03621E-02	4.56367E-03	4.19631E-03	1.63962E-03
Rosenbrock	$-10 < x_i < 10$	2.64842E-04	2.61154E-03	1.01552E-03	9.16315E-04	5.47507E-04
Rastringin	$-5.12 < x_i < 5.12$	8.96081E-04	9.97695E-01	6.21183E-02	3.58105E-03	2.23448E-01

Resultados

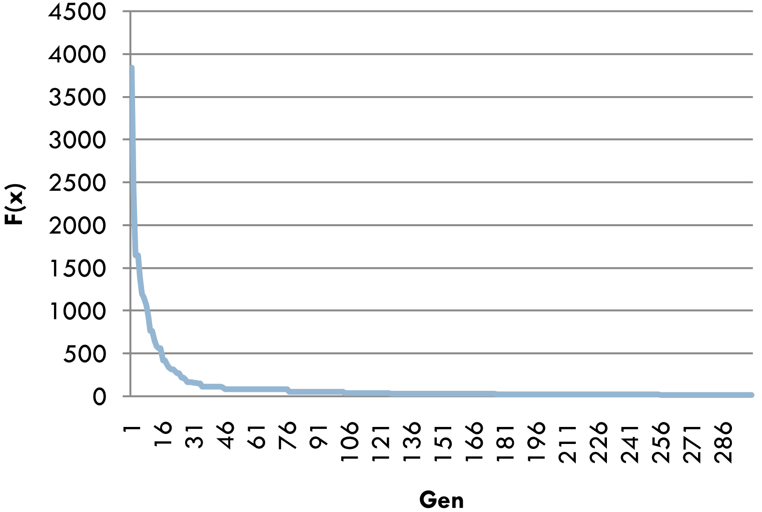
	Intervalo	Mejor	Peor	Media	Mediana	Desviación Estándar
DIM 10						
Sphere	$-600 < x_i < 600$	7.04279E-05	2.80888E-04	1.82826E-04	1.87302E-04	5.61840E-05
Griewank	$-600 < x_i < 600$	1.65654E-05	7.17427E-05	4.09305E-05	4.20921E-05	1.33490E-05
Ackley	$-10 < x_i < 10$	1.23249E-02	2.48941E-02	1.93264E-02	1.96231E-02	3.49406E-03
Rosenbrock	$-10 < x_i < 10$	3.42076E-02	1.15288E-01	6.04197E-02	5.26731E-02	2.18416E-02
Rastringin	$-5.12 < x_i < 5.12$	1.95447E-02	3.57272E+00	1.47668E+00	1.05266E+00	1.08807E+00

Resultados

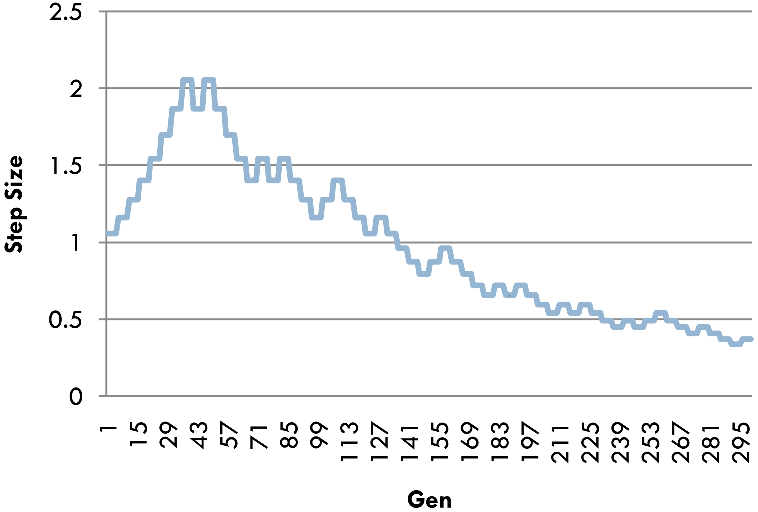
	Intervalo	Mejor	Peor	Media	Mediana	Desviación Estándar
DIM 20						
Sphere	$-600 < x_i < 600$	4.12598E-04	1.51723E-03	8.59681E-04	8.69425E-04	2.57094E-04
Griewank	$-600 < x_i < 600$	5.59018E-05	2.10015E-04	1.18548E-04	1.22605E-04	3.27922E-05
Ackley	$-10 < x_i < 10$	2.17126E-02	3.42767E-02	2.90389E-02	3.08381E-02	4.22667E-03
Rosenbrock	$-10 < x_i < 10$	3.16942E-01	3.09396E+00	1.13915E+00	8.28759E-01	7.57362E-01
Rastringin	$-5.12 < x_i < 5.12$	1.93206E-01	1.01444E+01	4.51284E+00	5.15190E+00	2.65504E+00

Rosenbrock (Dim = 20)

F(x) vs. Gen

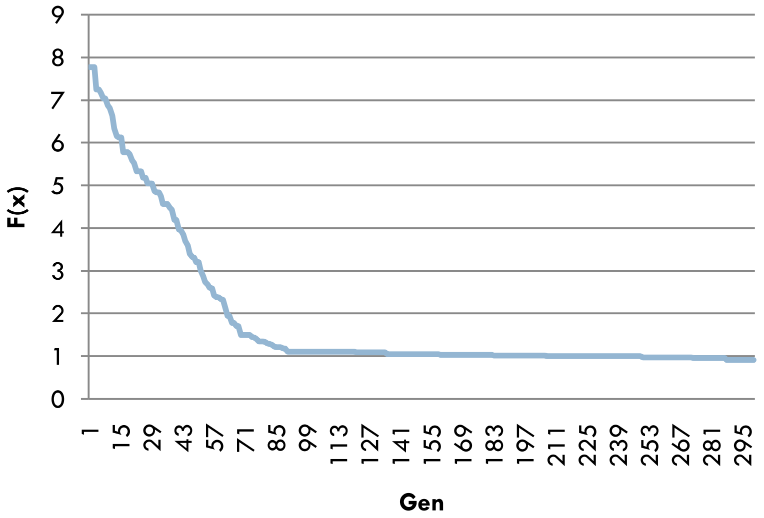


Step Size vs. Gen

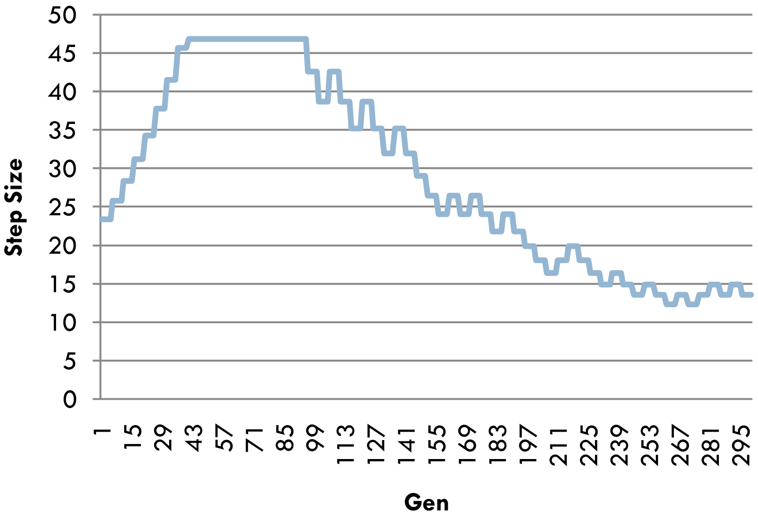


Grienwank (Dim = 20)

F(x) vs. Gen

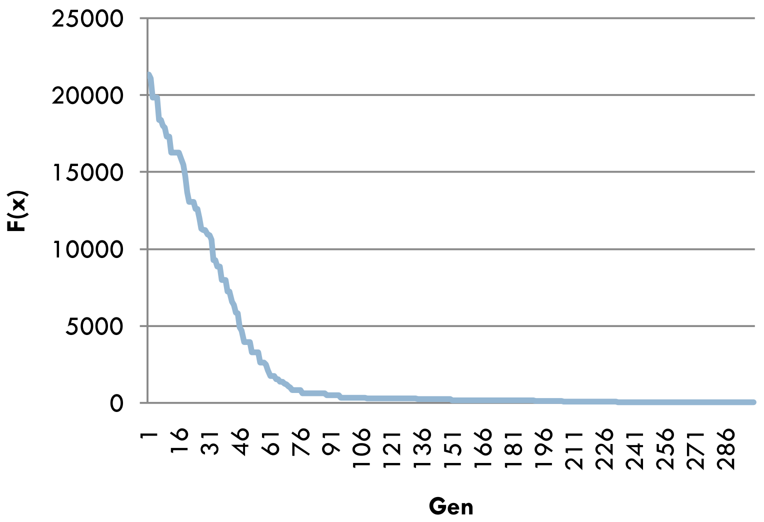


Step Size vs. Gen

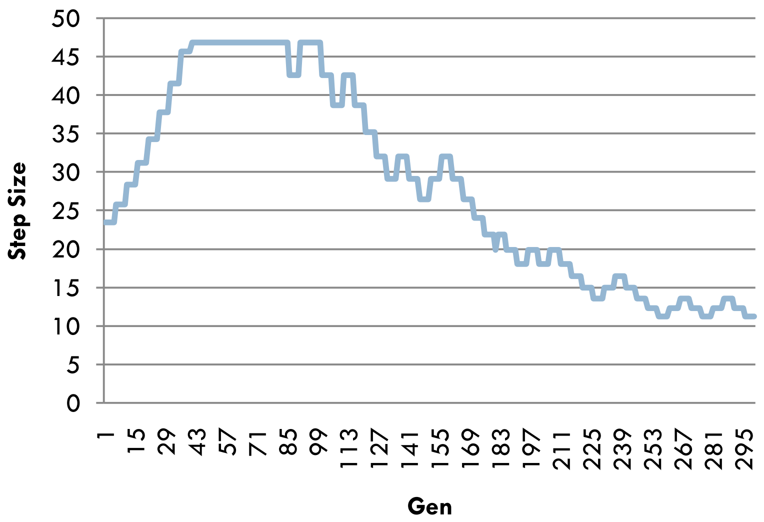


Sphere (Dim = 20)

F(x) vs. Gen

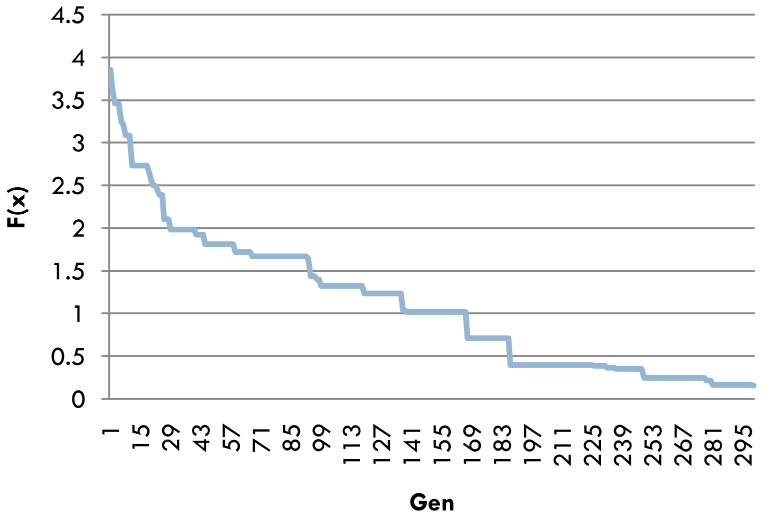


Step Size vs. Gen

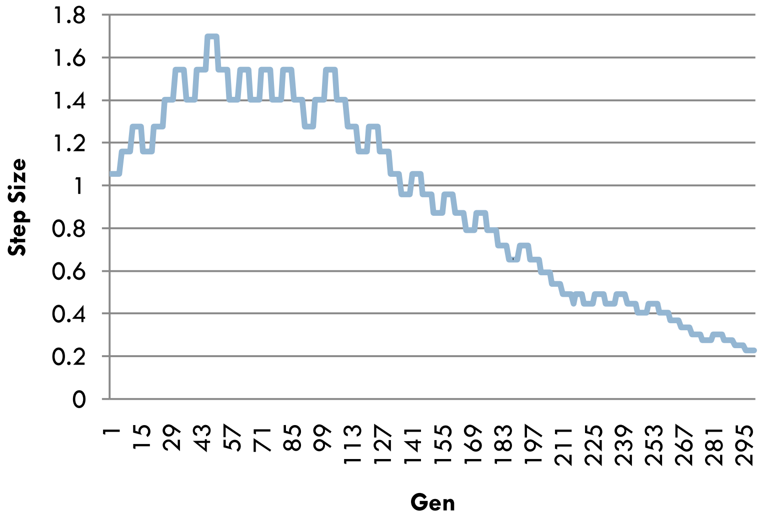


Ackley (Dim = 20)

F(x) vs. Gen

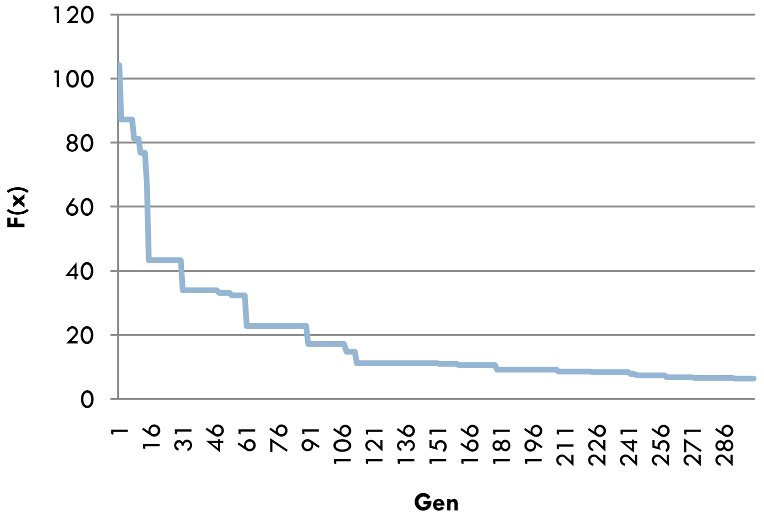


Step Size vs. Gen

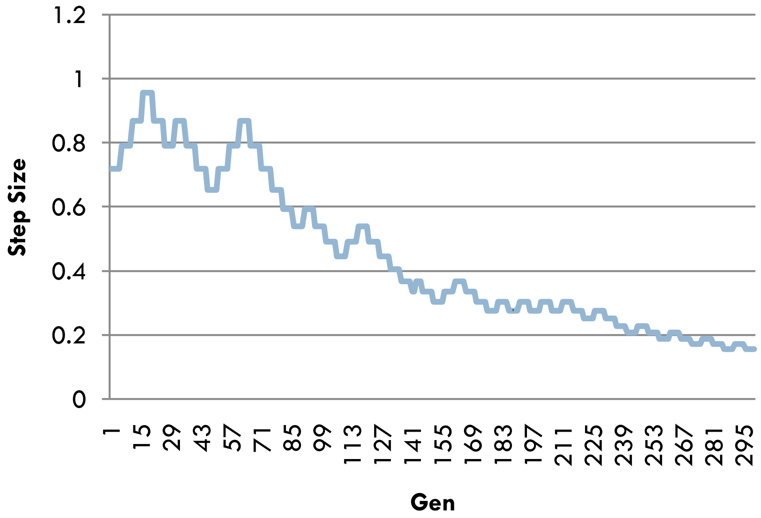


Rastringin (Dim = 20)

F(x) vs. Gen



Step Size vs. Gen



Conclusiones

- Los RRTs son una herramienta muy poderosa, debido a la rápida y buena exploración en el espacio de búsqueda. El problema de la exploración es uno de los más comunes al trabajar con los EDAs, pues en muchas ocasiones presentan una convergencia acelerada.

- Tomar la estructura de un RRT e incorporarla a un EDA puede llegar a ser muy ventajoso, pero hay que tener cuidado con adaptar la estructura del RRT y ajustarlo al problema que queremos resolver.

- Los RRTs por si mismos pueden ser insuficientes para resolver problemas específicos, pero al incorporarlo con otras técnicas puede a llegar a ser una herramienta de gran utilidad.

- Para un tamaño de paso pequeño, se puede obtener mayor precisión, pero al mismo tiempo, la matriz de covarianzas tiende a no ser positiva definida cuando la solución se acerca a un óptimo, haciendo que la factorización de Cholesky tenga problemas.

!GRACIAS¡