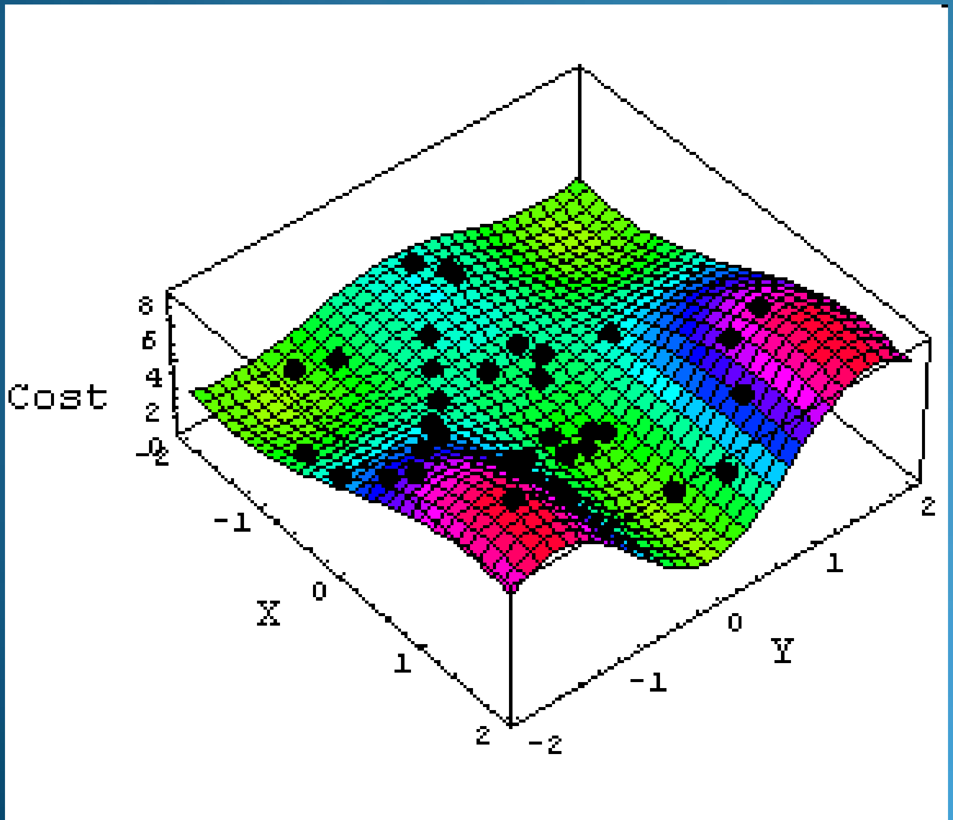
Optimización por medio de Algoritmos Genéticos y su convergencia por medio de cadenas de Markov



Carlos J. Zapata G
Carlos A. Ramirez V

Universidad Tecnológica de Pereira

Fundamentos Matemáticos de los Algoritmos Genéticos

Introducción

- Los AG provienen de la familia de modelos computacionales basados en la evolución

- Holland fue el primero quien le dio un formalismo y el nombre a los algoritmos genéticos

- Son modelos basados en una población inicial a los cuales se les aplica los operadores selección, recombinación y mutación

Ejemplo

- Se tiene N=6 producto max costo sujeto al volumen
- $Max=3x_1+2x_2+8x_3+4x_4+6x_5+12x_6$
- S.A $5x_1+4x_2+4x_3+2x_4+4x_5+6x_6 \leq 22$

- Población inical

0	0	1	1	0	1
0	1	1	1	1	0
0	0	0	1	1	1
1	1	1	1	0	0

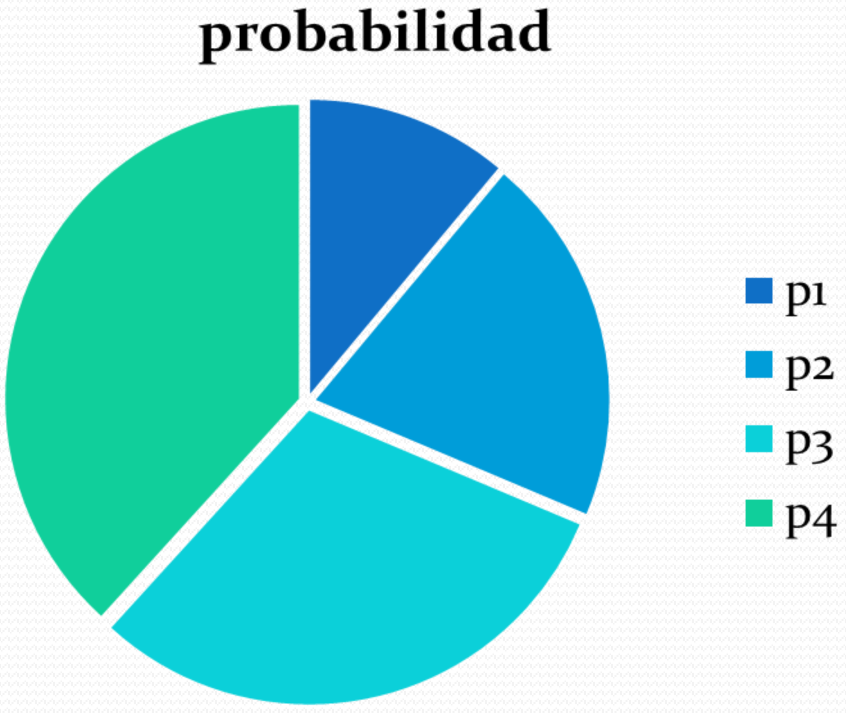
Ejemplo

- Función objetivo infactibilidad volumen
- Agente 1 =24 12
- Agente 2=20 14
- Agente 3=22 12
- Agente 4=17 15

- El Incumbente es el 1

Ejemplo

- En el operador selección se utiliza la ruleta
- P1=24/83=0,289
- P2=20/83=0,529
- P3=22/83=0,794
- P4=17/83=0,999



Ejemplo

- Derechos para los siguientes descendientes: se obtienen de la ruleta, en este caso supongamos que
- $p1=1$
- $P2=2$
- $P3=0$
- $P4=1$
- Y aleatoriamente se hace la recombinación p1 con p2 y p1 con p4

Recombinación

0	0	1	1	0	1
0	1	1	1	1	0

0	0	0	1	1	1
1	1	1	1	0	0

Cruce

- haciendo un cruce de un solo punto en las casillas 2 se obtiene para los dos casos

0	0	1	1	1	0
0	1	1	1	0	1

0	0	1	1	0	0
1	1	0	1	1	1

Mutación

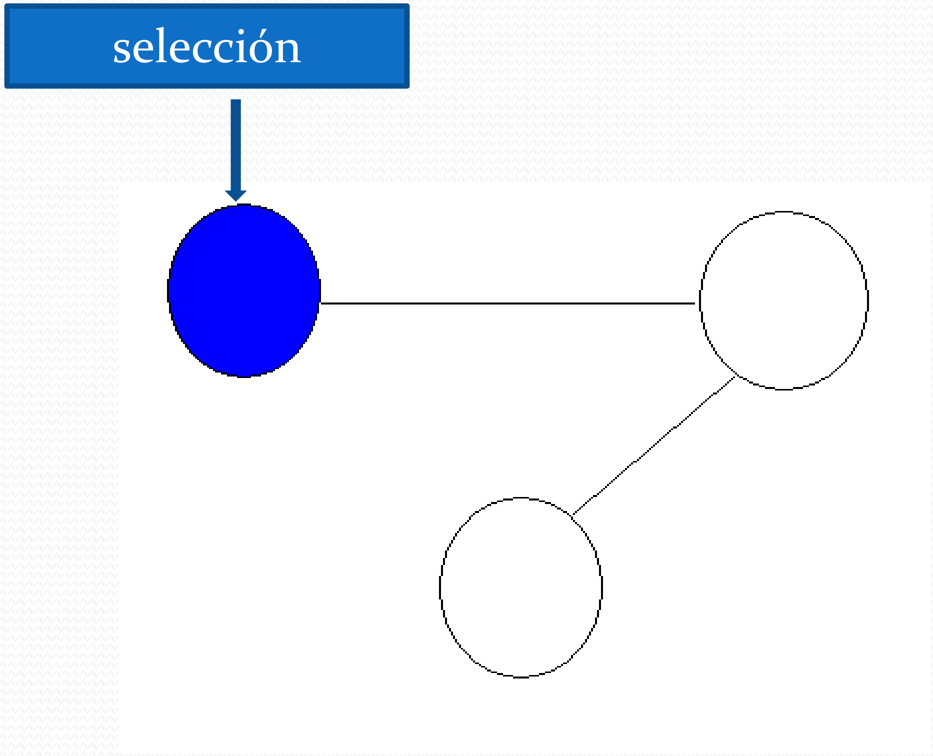
Función objetivo	infactibilidad volumen
P1=18	10
P2=26	16
P3=12	6
P4=27	21

- La mutación se hace sobre toda la población 6*4=24
- Tomemos el 5%, entonces 0,05*24=1,2 muta un gen de toda la población.

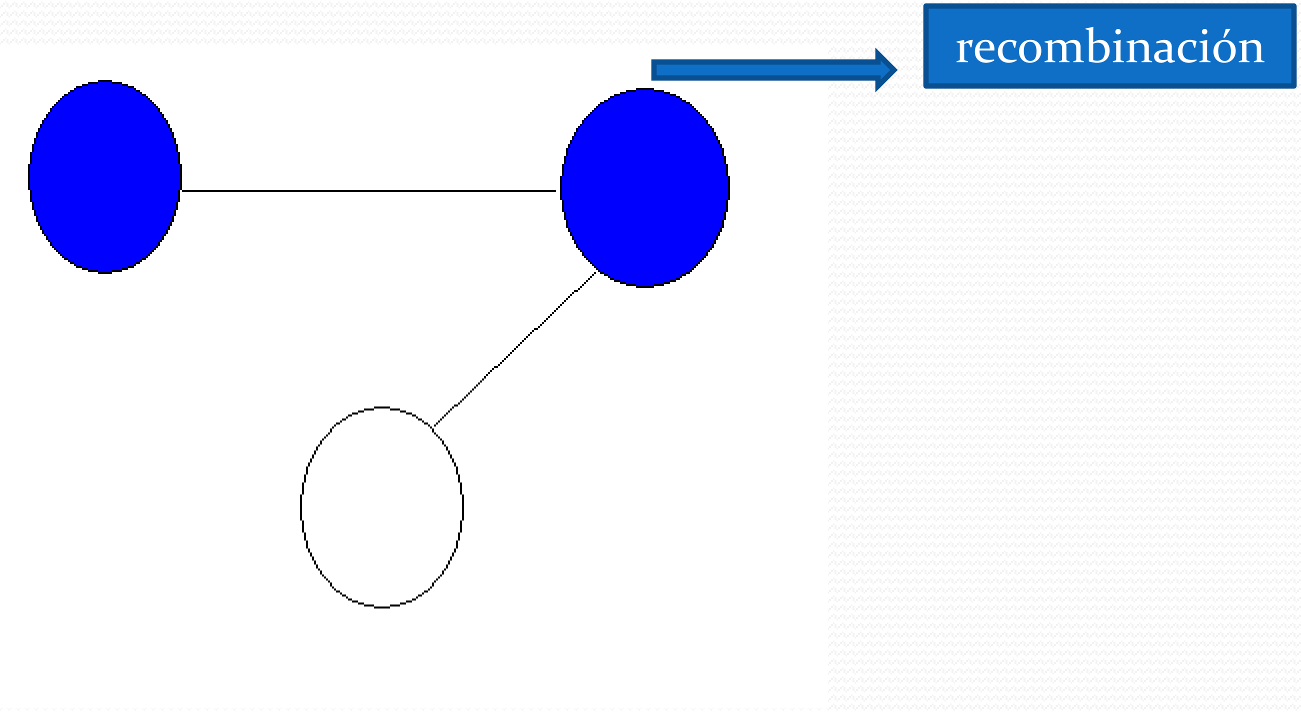
Operadores

- Crossover o recombinación: los nuevos individuos se obtienen luego de combinar a dos individuos de la generación anterior $I^2 \rightarrow \mathrm{I}$
- Mutación es un operador que actúa sobre toda la población randomicamente y lo hace entre el [1%-5%]

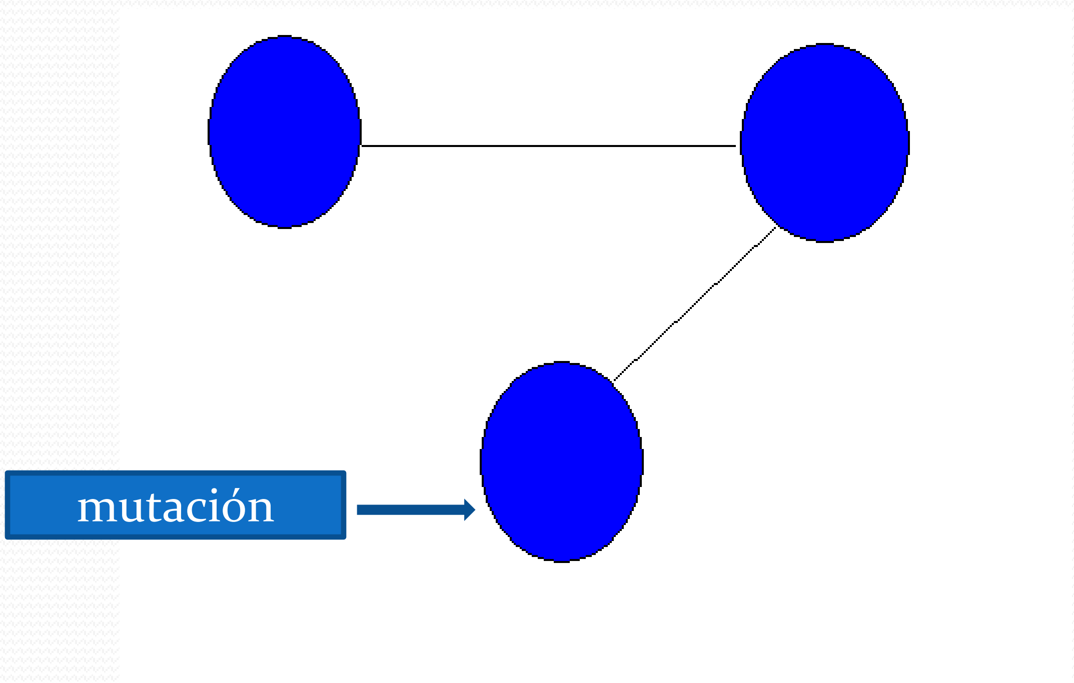
Algoritmo Genético



Algoritmo Genético



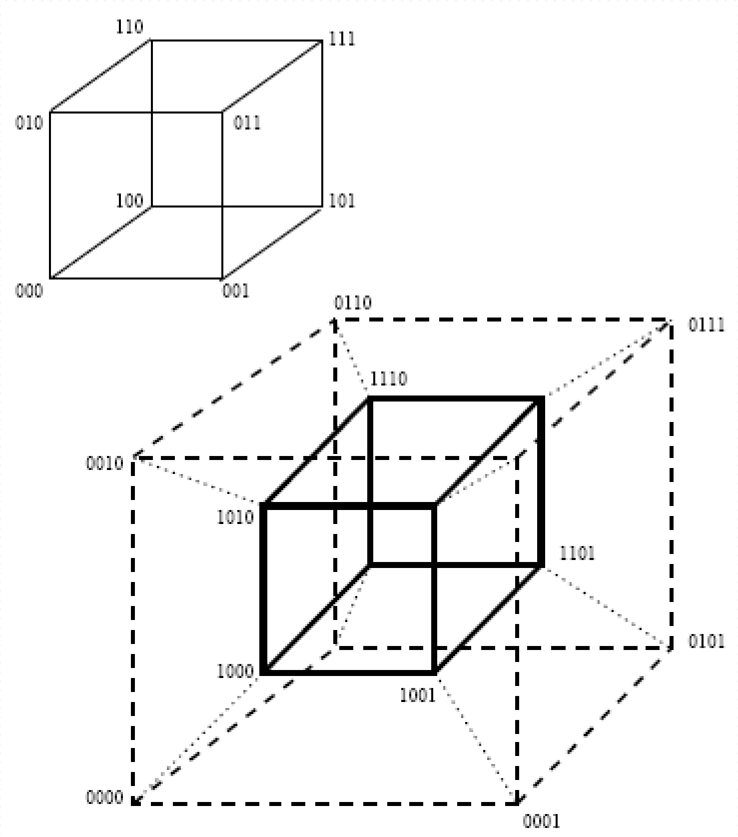
Algoritmo Genético



Teorema de los esquemas

- S={0,1} un alfabeto binario, para construir un esquema se necesita ampliar el alfabeto por medio del símbolo $*$
- Un esquema será cualquier ristra de caracteres formada a partir de S={0,1,*} si la ristra es l el número de esquemas existentes es 3^l ejemplo.
- La lista de longitud 4 y el esquema Q=(*,1,*,0) se empareja con los cromosomas (0,1,0,0),(0,1,1,0),(1,1,0,0),(1,1,1,0). Desde el punto de vista geométrico un esquema equivale a un hiperplano en el espacio de búsqueda

Geometría del esquema



Propiedades de los esquemas

- El orden del esquema O(Q) hace referencia al número de posiciones fijas. Ejemplo

Q_1=(*,0,1,*) O(Q_1) =2

Q_2 =(*,*,*,1) O(Q_2) =1

Q_3=(1,1,0,*) O(Q_3) =3

El concepto de orden del esquema se utiliza para calcular la probabilidad de supervivencia del esquema con relación al operador de mutación.

Propiedades de los esquemas

- Longitud del esquema $\delta\ (Q)$ se define como la distancia entre la primera y la ultima posición fija de la ristra. Ejemplo
- $\delta\ (Q_1)$=3-2=1
- $\delta\ (Q_2) = 4 - 4 = 0$
- $\delta\ (Q_3) = 3 - 1 = 2$
- El concepto de longitud se utiliza para calcular la probabilidad de supervivencia del esquema frente al operador de cruce

Propiedades de los esquemas

- ε(Q,t) denota el número de ristras que en una determinada población en el tiempo t se asocian con el esquema Q. ejemplo, sea la población $p_t = \{(0,1,0,0), (1,0,0,1), (1,1,1,1), (0,0,0,1)\}$ y los siguientes esquemas Q_1=$(*,*,0,1)$ y $Q_2(*,*,*,0)$ entonces $\varepsilon(Q_1$,t)=2 y $\varepsilon(Q_2$,t)=1
- Otra propiedad es la función de adaptación en el tiempo t y se denota por eval(Q,t) y se define como la media de la función objetivo de todas las ristras que en la población se asocian con el esquema Q

Teorema de los esquemas

- Sea $\underset{F(t)}{\longrightarrow}$ la media de al funcion objetivo extendida a toda la poblacion t. es decir $\underset{F(t)}{\longrightarrow}$ =F(t)/λ donde F(t) es la suma de las funciones objetivo de todos los individuos en al poblacione ne el tiempo t
- Teorema: $E_{sel,cru,mut}(\varepsilon(Q,t+1))\geq \varepsilon(Q,t)$ eval(Q,t)/ $\underset{F(t)}{\longrightarrow}$ [1-p_c $\frac{\delta(Q)}{l-1}$-O(Q)p_m]

Matrices

- Sea A_{n*n} una matriz
- Es no negativa si $a_{ij} \geq 0$ para todo i y j
- Es positiva si $a_{ij} > 0$ para todo i y j
- Sea B_{n*n} matriz no negativa
- Es estocástica si $\sum_{j=1}^{n} b_{ij}$=1
- Es primitiva si existe k entero positivo tal que $B^k > 0$
- Es irreducible si para todo b_{ij} existe k entero tal que ${b_{ij}}^k > 0$
- Reducible si no es irreducible

Matrices

- Sea C_{n*n} una matriz estocástica
- Diagonal positiva si cada elemento en su diagonal es positivo
- Columna permisible si al menos tiene una componente positiva en cada fila
- Es estable si tiene filas idénticas
- Lema B es una matriz permisible si y solo si se puede llevar a la forma $\begin{bmatrix} C & 0 \\ R & T \end{bmatrix}$ C y T son cuadradas

Cadenas de Markov

- Un proceso de Markov es un proceso estocástico tal que la probabilidad de que la variable aleatoria X_n en el estado i_n solo dependa de la variable aleatoria anterior.
- P(X_n=i_n / X_0=i_0, …. . , X_{n-1}=i_{n-1})=

 P(X_n=i_n / X_{n-1}=i_{n-1})

La probabilidad de transición de n pasos para un proceso de Markov es definida como

P_{ij}(n)=P(X_n=j/X_0=i)=P($X_{n+k} = j/X_n$=i)

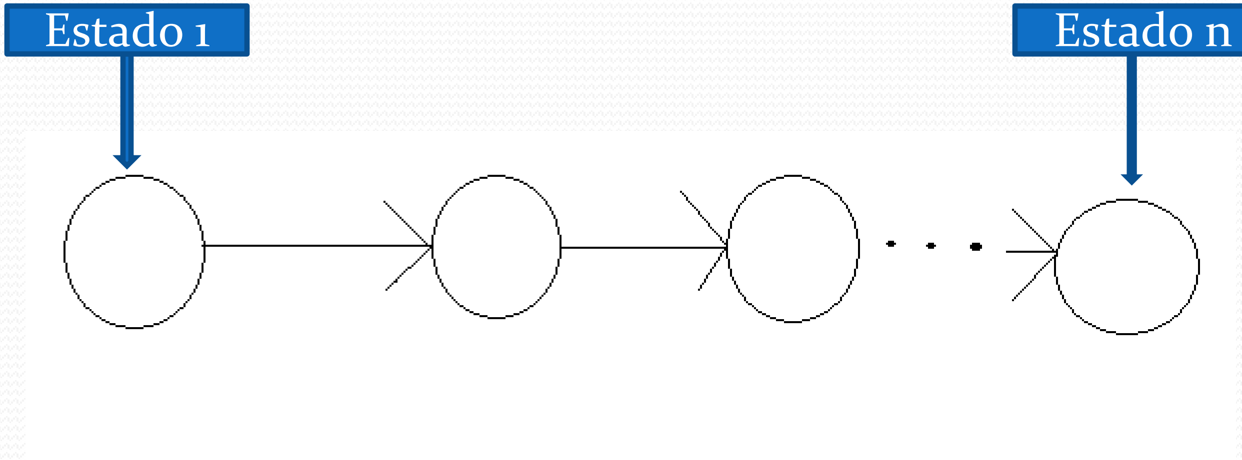
Cadenas de Markov

- Definicion Matriz de transición de una cadena de markov es la matriz $\mathbb{p}$ con entradas p_{ij}
- Definición el vector de probabilidad inicial es el vector π_0={p(X_0=j)}
- Nota observe que la matriz de transición es una matriz estocástica
- El vector de probabilidad inicial y la matriz de transición determinan completamente la cadena de markov

Estados de la cadena

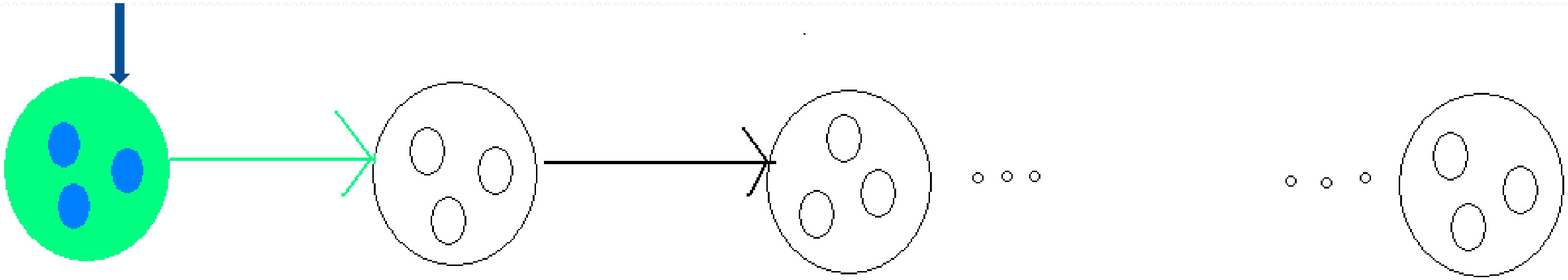
- Se clasifican los estados con base a: si es posible ir de un estado dado i a otro estado dado j, esto es si existe un paso n tal que la probabilidad de ir de i a j es mayor que P_{ij}>0

Cadena de Markov



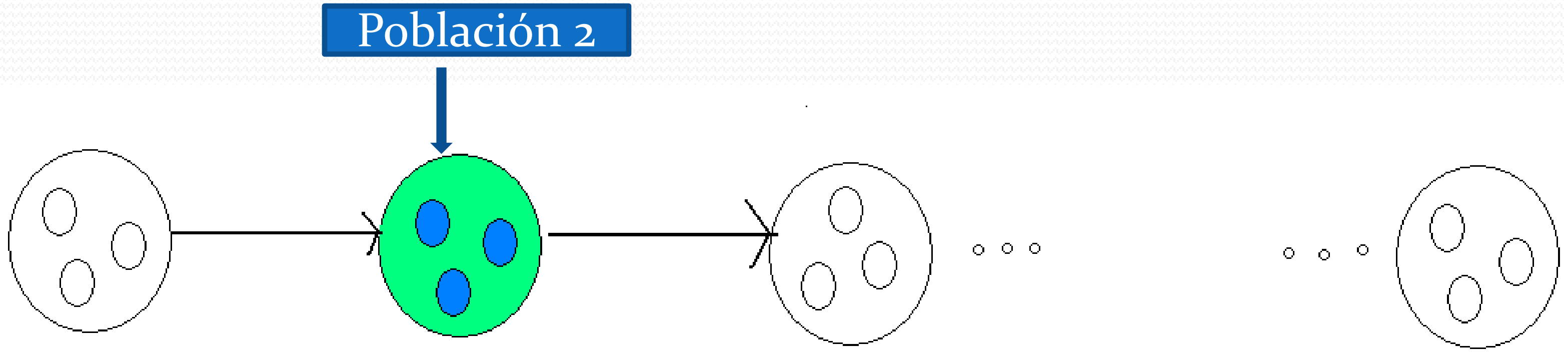
Cadenas de Markov

Población 1



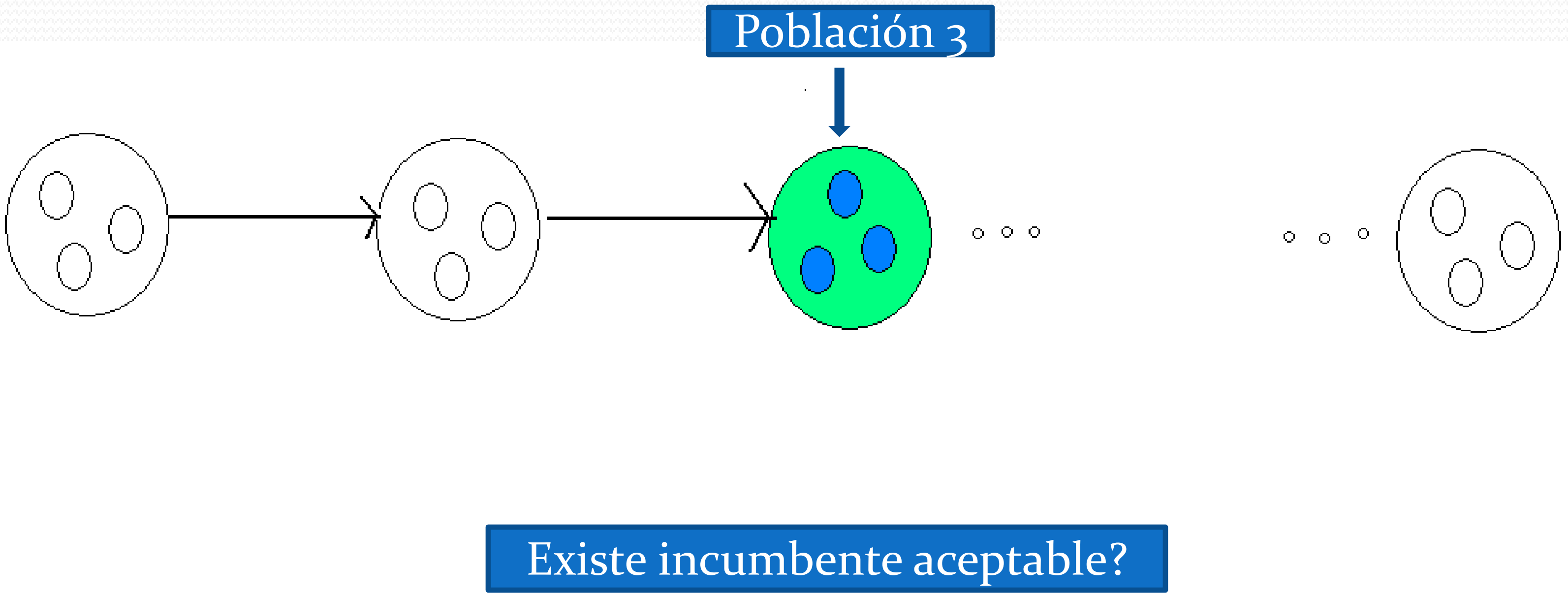
Existe incumbente aceptable?

Cadenas de Markov

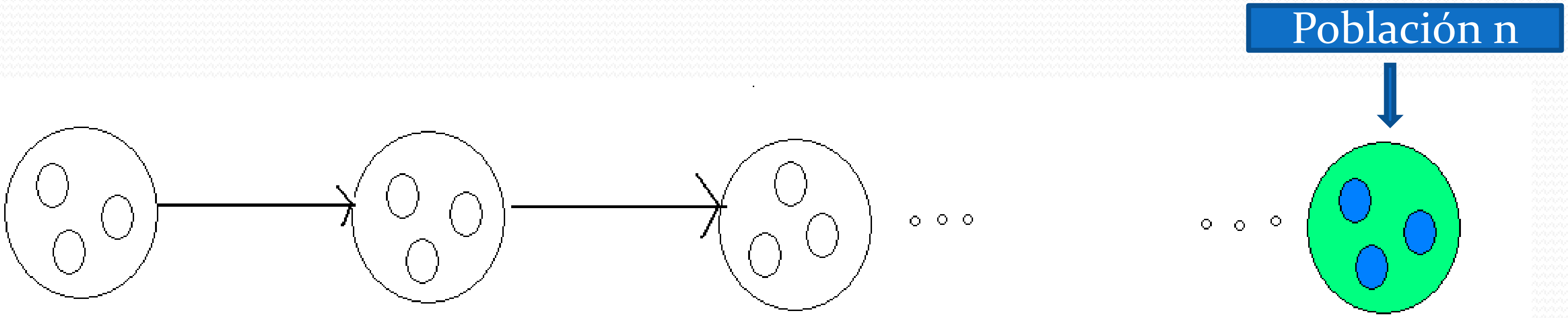


Existe incumbente aceptable?

Cadenas de Markov



Cadenas de Markov



Relación de orden

- Sea U un conjunto no vacío y R una relación entre dos elementos de U, entonces aRb es el hecho de que a esta en relación con b.
- Ejemplos
- R es reflexiva si xRx para todo $x \epsilon U$
- R es simétrica si xRy, entonces yRx para todo $x,y \epsilon U$
- R es transitiva si xRy e yRz, entonces xRz para todo $x,y,z \epsilon U$
- R es una relación de equivalencia si R es reflexiva, simétrica y transitiva.

Relación de orden

- Un elemento a de U es llamado minimal si aTx implica xTa para todo $x \epsilon U$
- Definición: el elemento minimal es llamado conjunto ergodico, los elementos restantes son llamados conjuntos trascendentes.
- Si un conjunto ergodico contiene un solo elemento se le llamara estado absorbente
- Teorema: un estado i es absorbente $\Leftrightarrow P_{ii} = 1$
- Nota: la matriz de transición es estocástica entonces las entradas de la fila de un estado absorbente son cero excepto $P_{ii} = 1$.

Demostración

- Teorema: para cualquier cadena de Markov finita, no importa el estado donde inicie el proceso, la probabilidad después de n pasos de que esté en un estado ergodico tiende a 1 cuando n$\rightarrow \infty$
- Proposición: $\mathbb{P}=\begin{bmatrix} S & 0 \\ R & Q \end{bmatrix}$ donde S representa la cadena cuando ha alcanzado un estado ergodico; O la matriz con entradas cero; R representa la transición de estados transitorios a ergodicos y Q representa la cadena mientras esta en estados transitorios.

Demostración

-
- Pero $\mathbb{p} = \begin{bmatrix} I & 0 \\ R & Q \end{bmatrix}$ ademas Q^k →o cuando k → ∞

- Teorema la matriz (I- Q) tiene inversa y
- $(I - Q)^{-1}$=I+Q+$Q^2 + \cdots = \sum_{k=0}^{\infty} Q^k$

- Definicion para una cadena absorbente, se define N=$(I - Q)^{-1}$ donde N se llama matriz fundamental

Demostración

- Teorema sea $\mathbb{p} = \begin{bmatrix} I & 0 \\ R & Q \end{bmatrix}$ la forma canonica de la matriz de transicion de uan cadena de markov, entonces $\mathbb{p}^k$ converge a la matriz $\mathbb{p}^\infty = \begin{bmatrix} S^\infty & O \\ R_\infty & O \end{bmatrix}$ cuando k $\rightarrow \infty$ donde S^∞ $y\ R_\infty$ denotan S^k y NR S^k

Cuando k $\rightarrow \infty$

Definicion:

$$C^* M^* S^* = \begin{pmatrix} C & \cdots & \\ \vdots & \ddots & \vdots \\ & \cdots & C \end{pmatrix} \begin{pmatrix} M & \cdots & \\ \vdots & \ddots & \vdots \\ & \cdots & M \end{pmatrix} \begin{pmatrix} S & \cdots & \\ \vdots & \ddots & \vdots \\ & \cdots & S \end{pmatrix}$$

Demostración

-
- $C^*M^*S^* = \begin{bmatrix} \mathbb{p} & \cdots & \\ \vdots & \ddots & \vdots \\ & \cdots & \mathbb{p} \end{bmatrix}$
- Sea E la matriz del operador elitista, lo que hace es verificar si un individuo contiene, o no un invividuo mas apto que el super individuo, si es asi lo reemplaza
- $e_{ij} = \begin{cases} 1, \ si \ \ F(\pi_0(i)) < f(b) = f((\pi_0(j))) \\ \quad 0, \ en\ otro\ caso \end{cases}$

Demostración

- Note que solo existe un 1 por fila en tonce s E se puede escribir E= $\begin{bmatrix} E_{11} & \cdots & \\ \vdots & \ddots & \vdots \\ E_{2^l 1} & \cdots & E_{2^l 2^l} \end{bmatrix}$
- E_{11}=I
- $\mathbb{p}^*$= $\begin{bmatrix} \mathbb{p} & \cdots & \\ \vdots & \ddots & \vdots \\ & \cdots & \mathbb{p} \end{bmatrix} \begin{bmatrix} E_{11} & \cdots & \\ \vdots & \ddots & \vdots \\ E_{2^l 1} & \cdots & E_{2^l 2^l} \end{bmatrix}$

Demostración

-
- $\mathbb{p}^*=\begin{bmatrix} \mathbb{p}E_{11} & \cdots & \\ \vdots & \ddots & \vdots \\ \mathbb{p}E_{2^l 1} & \cdots & \mathbb{p}E_{2^l 2^l} \end{bmatrix}$
- R= $\begin{matrix} \mathbb{p}E_{11} \\ . \\ \mathbb{p}E_{2^l 1} \end{matrix}$ y T=$\begin{bmatrix} \mathbb{p}E_{22} & \cdots & \\ \vdots & \ddots & \vdots \\ \mathbb{p}E_{2^l 2} & \cdots & \mathbb{p}E_{2^l 2^l} \end{bmatrix}$
- Y como $\mathbb{p}E_{11} = \mathbb{p}$ entonces
- $\mathbb{p}^*=\begin{bmatrix} \mathbb{p} & O \\ R & T \end{bmatrix}$ siendo así reducible

Preguntas

Bibliografía

- Carlos,J. Zapata;Análisis probabilísticos y simulación utp 2011.
- Ramon,G. Antonio,H. Eliana,M Tecnicas metaheuristicas de optimizacion utp 2008
- Back,T; Evolutionary Algorithm in Theory and Practice:Oxford University Press 1996
- Gunter,R.Alexandru,A :Convergence Properties of Some Multi-Objetive Evolutionary Algorithms