VC4 — Descenso del Gradiente 03MIAR – Algoritmos de Optimización

VIU Universidad Internacional de Valencia



El Problema de Asignación Cuadrática (QAP)

JCC 2016, Rosario

por

&

López Restrepo, Jheison jalopez@ungs.edu.ar

Investigador/Docente UNGS Maestria en Ingenieria - UdeA Rojas Paredes, Andrés

arojas@ungs.edu.ar
Investigador/Docente de UNGS
Doctorando en FCEyN-UBA









Agenda

- Revisión de la Complejidad. Cálculo de operaciones.
- Descenso del Gradiente
- Problemas del Trabajo Práctico



 $Pg.: \langle N^o \rangle$

VC4 — Complejidad 03MIAR – Algoritmos de Optimización

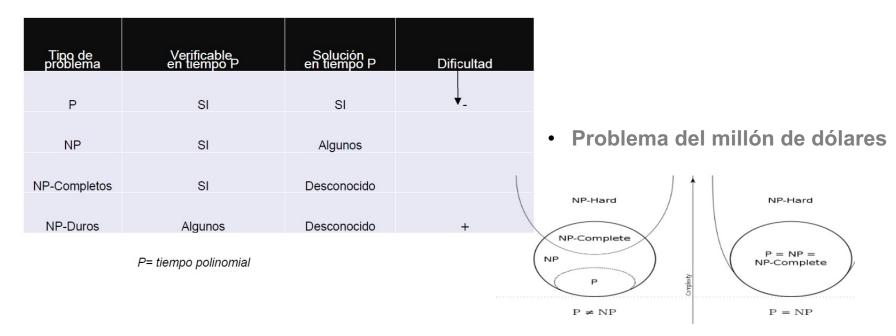
Agenda

- Revisión de la Complejidad. Cálculo de operaciones.
- Descenso del Gradiente
- Problemas del Trabajo Práctico



Introducción(Simplificada) a la teoría de la complejidad

• ¿P=NP? O, ¿Qué problemas se pueden resolver fácilmente con computación?





NP: no se conoce solución en tiempo polinomial

NP-Completo: si un problema NP-Completo se demuestra que es P entonces P=NP! (El resto de NP se pueden transformar con complejidad Polinomial) NP-Duros: Al menos tan complejo como cualquier NP pero puede que no sea NP

Introducción(Simplificada) a la teoría de la complejidad

Ejemplos:

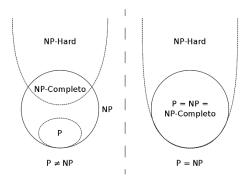
- NP-Completo : Satisfacibilidad booleana(SAT)(1971-Stephen Cook).

* Todos los NP-Completos se "reducen" a SAT

https://es.wikipedia.org/wiki/Anexo:Problemas_NP-completos

- NP-Duro

* Problema del Agente Viajero(TSP)





Introducción(Simplificada) a la teoría de la complejidad

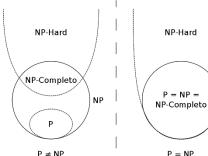
;P=NP?

Resolverlo supondría encontrar "atajos" para resolver problemas NP a partir de problemas P (reducción de un problema a otro)

Todo indica que P no es igual a NP por lo que "casi" asumimos que no podemos abordar estos problemas computacionalmente(*) con algoritmos exactos!

Es muy probable que solo podamos resolver estos problemas por métodos aproximados o usando heurísticas.







http://www.claymath.org/millennium-problems/p-vs-np-problem

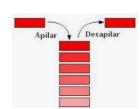
Análisis de Algoritmos

Contar operaciones elementales

- o Suma, resta, multiplicación y división
- Asignaciones
- o Condiciones , llamadas a funciones externas y retornos
- Accesos a estructuras de datos
- o Ojo con los bucles anidados! 2 bucles $\Rightarrow n^2$
- Orden de complejidad
 - ✓ Decimos que un algoritmo tiene orden de complejidad o(f) si su tiempo de ejecución en operaciones elementales está acotado, salvo por una constante, por f para todos los tamaños de entrada a partir de un cierto n_0



a+=b	a = a + b
a-=b	a = a - b
a*=b	a = a * b
a/=b	a = a / b
-0/ h	0/ h



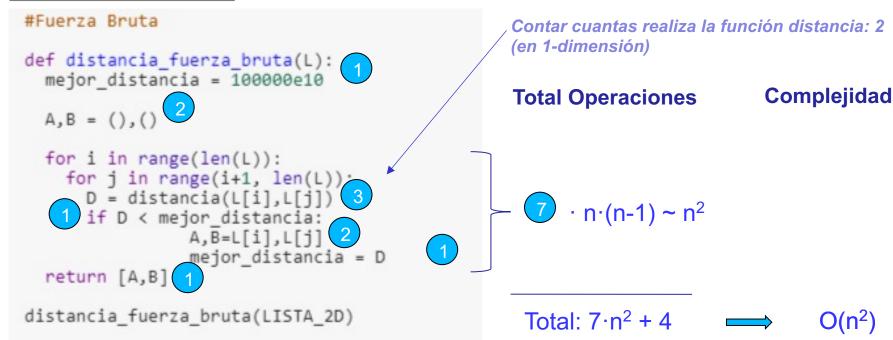






Análisis de Algoritmos. Sin recursividad

Contar operaciones elementales.



Análisis de Algoritmos. Ecuaciones en recurrencia

El termino n-simo depende de los términos anteriores T(n) = T(n-1) + T(n-2)

```
#Sucesión de Fibonacci
#https://es.wikipedia.org/wiki/Sucesi%C3%B3n de Fibonacci
#Calculo del termino n-simo de la suscesión de Fibonacci
def Fibonacci(N:int):
  if N < 2:
    return 1
  else:
    return Fibonacci(N-1)+Fibonacci(N-2)
Fibonacci(5)
8
```

$$T(n) - T(n-1) - T(n-2) = 0$$

Eçuación característica:

$$x^2 - x - 1 = 0$$

$$x_1=rac{1+\sqrt{5}}{2} \qquad x_2=rac{1-\sqrt{5}}{2}$$

La solución es

$$T(n) = A \cdot \left(\frac{1+\sqrt{5}}{2}\right)^n + B \cdot \left(\frac{1-\sqrt{5}}{2}\right)^n$$

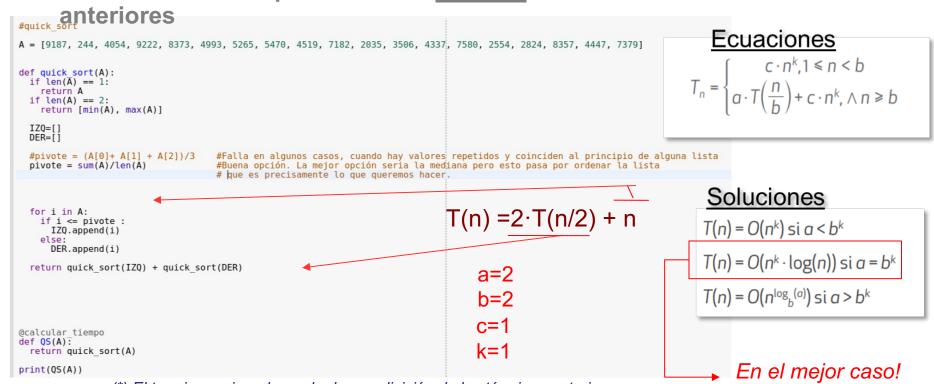


A y B dependen de los términos iniciales

 $Pg.: \langle N^o \rangle$

Análisis de Algoritmos. Ecuaciones en recurrencia

El termino n-simo depende de una división de los términos



(*) El termino n-simo depende de una división de los términos anteriores

Análisis de la complejidad temporal

Ordenes de complejidad (de menor a mayor)

O(1)	Orden constante
O(log n)	Orden logarítmico
O(n)	Orden lineal
O(n log n)	Orden cuasi-lineal
O(n²)	Orden cuadrático
O(n³)	Orden cúbico
O(nª)	Orden polinómico
O(2 ⁿ)	Orden exponencial
O(n!)	Orden factorial



















Tabla 1

Comparación de los órdenes de complejidad según el tamaño del problema

n	logn	n²	2^n	n!
2	1	4	4	2
8	3	64	256	40.320
32	5	1.024	4,3 × 10 ⁹	2,6 × 10 ³⁵
100	6	10 ⁴	1,2 ×10 ²⁷	9,3 × 10 ¹⁷⁷



Sebar

VC4 — Descenso del Gradiente 03MIAR – Algoritmos de Optimización

Agenda

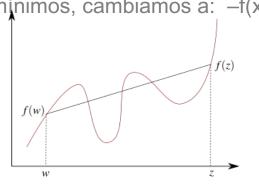
- Revisión de la Complejidad. Cálculo de operaciones.
- Descenso del Gradiente
- Problemas del Trabajo Práctico

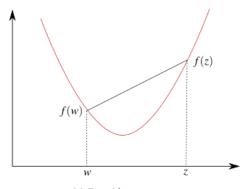


• Es un algoritmo **iterativo** de optimización para encontrar valores mínimos para funciones multi-varibles **convexas** y **diferenciables**(y por tanto continuas).

- Convexidad
 - Permitir asegurar que la óptimo que encontremos es global frente a que solo sea local
 - ¿El segmento que une dos puntos está siempre por encima de la función?

Si es cóncava y buscamos mínimos, cambiamos a: -f(x)





importante ·



 $Pg.: \langle N^o \rangle$

a) Función no convexa

b) Función convexa

- Diferenciable: derivable en n variables(dimensiones)
- **Gradiente**: Es un vector(diferente para cada punto = campo vectorial) cuyas coordenadas son las derivadas parciales:

$$oldsymbol{
abla} f(\mathbf{r}) = \left(rac{\partial f(\mathbf{r})}{\partial x_1}, \ldots, rac{\partial f(\mathbf{r})}{\partial x_n}
ight)$$

Ejemplo: Dada la función $f(x,y) = x^2 + 2x + y^2 + y^3 + xy$ el gradiente será:

$$abla f = \left(rac{\partial f}{\partial x},rac{\partial f}{\partial y}
ight) \qquad \left(2x+y+2,2y+3y^2+x
ight)$$

En el punto (x,y)

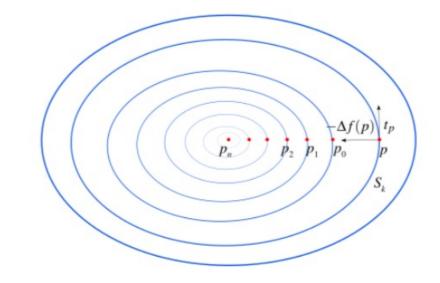


El procedimiento parte de un punto **p como solución aproximada(inicial)** y da pasos en el sentido opuesto(si minimizamos) al gradiente de la función en dicho punto.

$$p_{t+1} = p_t - \alpha_t \cdot \triangle f(p_t)$$

La elección de α, estará condicionada para que :

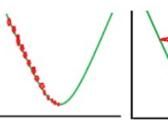
- p_{t+1} sea solución factible
- mejore el valor de f respecto a p_t : $f(p_{t+1}) \ge f(p_t)$

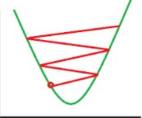


La elección del parámetro α_t , llamado tasa de aprendizaje (learning rate), es importante para hacer efectivo el proceso de acercamiento a la solución óptima.

- Un valor demasiado pequeño puede provocar exceso de iteraciones(Figura 1)
- Un valor demasiado alto puede provocar que el proceso no se acerque los suficiente(Figura 2)

$$p_{t+1} = p_t - \alpha_t \cdot \triangle f(p_t)$$





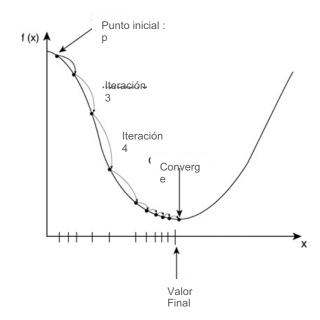
(Figure 1)

(Figure 2)

En general es buena estrategia ir reduciendo el valor de α_t dinámicamente a medida que nos aproximamos a la solución

- ¿Como sabemos que nos acercamos?:
 - la magnitud del gradiente
 - cantidad de iteraciones que hemos realizado

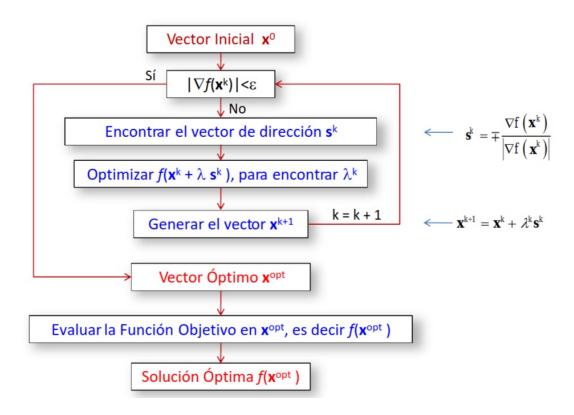
$$p_{t+1} = p_t - \alpha_t \cdot \triangle f(p_t)$$





http://www.bdhammel.com/learning-rates/

Diagrama de resolución





- Aprendizaje supervisado y la regresión lineal
 - ✓ En problemas de clasificación tratamos de clasificar los puntos en sus categorías correspondientes de tal manera que se minimice el error.
 - ✓ Para la medida de este error solemos utilizar el error cuadrático medio(regresión lineal) pero es posible usar otras medidas(por ejemplo: distancia Euclidea, Manhattan, ...)

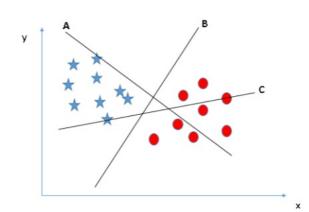
✓ La función que acumula los errores para los valores conocidos la llamamos función de coste. En el caso de la regresión línea es cuadrática, por tanto diferenciable y podemos obtener el gradiente.

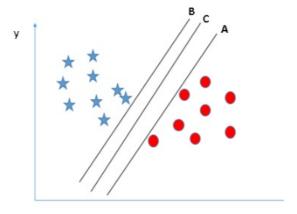


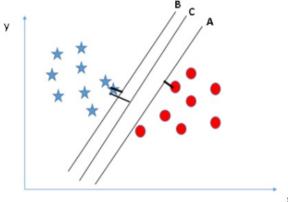
-20 -10 10 20 30 40 50 60

· Aprendizaje supervisado y la regresión lineal

¿Cómo identificar la mejor recta (hiperplano en N-dimensión) que separa los conjuntos de datos?





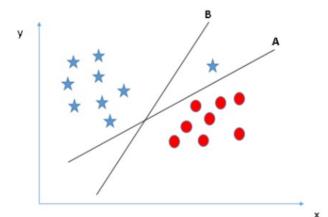


Seleccionar el que esté a mayor distancia de ambos conjuntos



Aprendizaje supervisado y la regresión lineal

¿Cómo identificar la mejor recta (hiperplano en N-dimensión) que separa los conjuntos de datos?



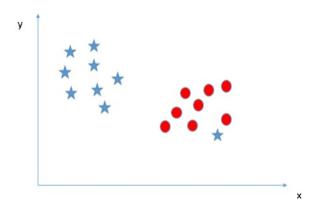
- 1°. Seleccionar el que mejor clasifique
- 2º . Seleccionar el que está a mas distancia



 $Pg.: \langle N^o \rangle$

Aprendizaje supervisado y la regresión lineal

¿Cómo identificar la mejor recta (hiperplano en N-dimensión) que separa los conjuntos de datos?



No siempre es posible clasificar 100%

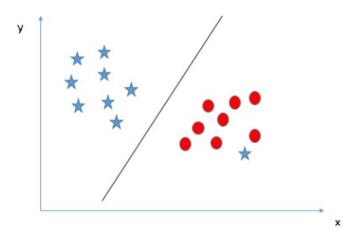
¿valores atípicos?



Pg.: ⟨N⁰⟩

Aprendizaje supervisado y la regresión lineal

¿Cómo identificar la mejor recta (hiperplano en N-dimensión) que separa los conjuntos de datos?



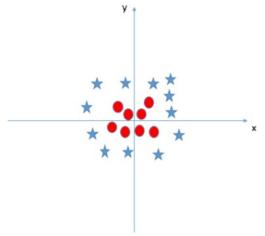
• Ignorar valores atípicos



 $Pg.: \langle N^o \rangle$

Aprendizaje supervisado y la regresión lineal

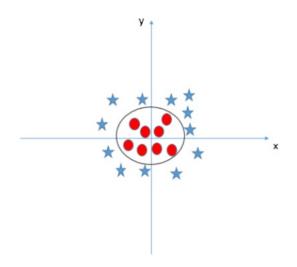
¿Cómo identificar la mejor recta (hiperplano en N-dimensión) que separa los conjuntos de datos?



- No parecen valores atípicos
- · Parece que si hay una clasificación
- Es imposible con hiperplanos lineales

Aprendizaje supervisado y la regresión lineal

Podemos extender el procedimiento a funciones no lineales!



Una función como $f(x,y) = x^2 + y^2$ lo resuelve

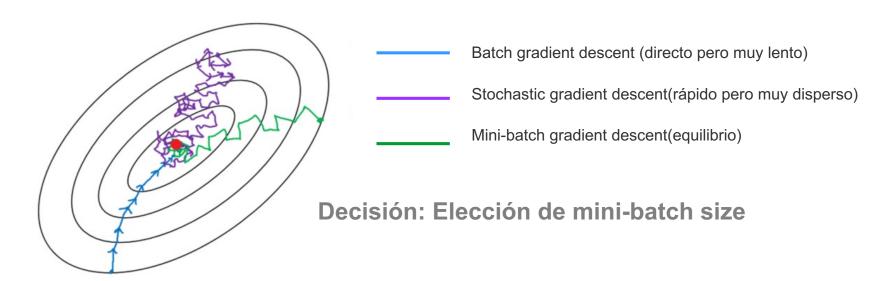


- Otra opción para función de coste(no cuadrática)
- Entropía cruzada(cross-entropy)
 - Para problemas con soluciones binarias
 - Definición: Número de bits diferentes. Mide como de diferentes son dos elementos.
 - Usada en algoritmos de clasificación de imágenes

- Dependiendo del Volumen de Datos
 - ✓ Descenso del gradiente por lotes (batch gradient descent)
 Calcula la desviación para todos los puntos en cada iteración!!!
- ✓ Descenso del gradiente estocástico(stochastic gradient descent)
 Calcula la desviación para un punto en cada iteración!!!
 - ✓ Descenso del gradiente por lotes reducido(mini-batch gradient descent)
 Mezcla de ambos conceptos



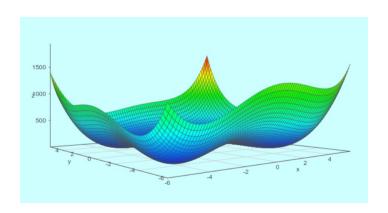
Dependiendo del Volumen de Datos





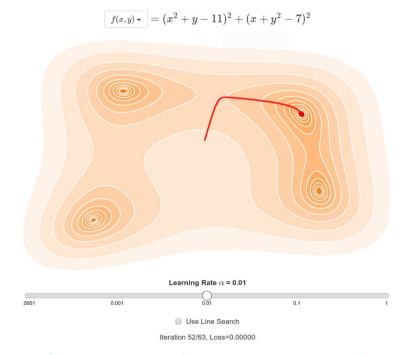
Pg.: ⟨N⁰⟩

Visualización



$$(x^2+y-11)^2+(x+y^2-7)^2$$

http://al-roomi.org/3DPlot/index.html



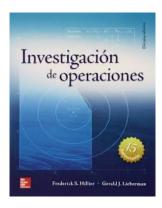


https://www.benfrederickson.com/numerical-optimization/

Ampliación de conocimientos

¿Qué es el Descenso del Gradiente? Algoritmo de Inteligencia Artificial | DotCSV https://www.youtube.com/watch?v=A6FiCDoz8_4

Hillier, F. S., y Lieberman, G. J. Investigación de Operaciones. Ciudad de México





VC4 — Problemas del Trabajo Práctico 03MIAR – Algoritmos de optimización

Agenda

- Revisión de la Complejidad. Cálculo de operaciones.
- Descenso del Gradiente
- Problemas del Trabajo Práctico



Problema 1. Organizar sesiones de doblaje(I)

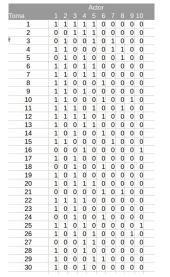
 Se precisa coordinar el doblaje de una película. Los actores del doblaje deben coincidir en las tomas en las que sus personajes aparecen juntos en las diferentes tomas. Los actores de doblaje cobran todos la misma cantidad por cada día que deben desplazarse hasta el estudio de grabación independientemente del número de tomas que se graben. No es posible grabar más de 6 tomas por día. El objetivo es planificar las sesiones por día de manera que el gasto por los servicios de los actores de doblaje sea el menor posible. Los datos son:

Número de actores: 10 Número de tomas : 30

Actores/Tomas: https://bit.ly/36D8luK

- 1 indica que el actor participa en la toma

- 0 en caso contrario







Problema 2. Organizar los horarios de partidos de La Liga(I)

- Desde la La Liga de fútbol profesional se pretende organizar los horarios de los partidos de liga de cada jornada. Se conocen algunos datos que nos deben llevar a diseñar un algoritmo que realice la asignación de los partidos a los horarios de forma que maximice la audiencia.
- Los horarios disponibles se conocen a priori y son los siguientes:

Viernes	20
Sábado	12,16,18,20
Domingo	12,16,18,20
Lunes	20





Problema 2. Organizar los horarios de partidos de La Liga(II)

- En primer lugar se clasifican los equipos en tres categorías según el numero de seguidores(que tiene relación directa con la audiencia). Hay 3 equipos en la categoría A, 11 equipos de categoría B y 6 equipos de categoría C.
- Se conoce estadísticamente la audiencia que genera cada partido según los equipos que se enfrentan y en horario de sábado a las 20h (el mejor en todos los casos)

	Categoría A	Categoría B	Categoría C
Categoría A	2 Millones	1,3 Millones	1 Millones
Categoría B		0.9 Millones	0.75 Millones





Problema 2. Organizar los horarios de partidos de La Liga(III)

- Si el horario del partido no se realiza a las 20 horas del sábado se sabe que se reduce según los coeficientes de la siguiente tabla
- Debemos asignar obligatoriamente siempre un partido el viernes y un partido el lunes

	Viernes	Sábado	Domingo	Lunes
12h	-	0.55	0.45	-
16h	-	0.7	0.75	-
18h	-	0.8	0.85	-
20h	0.4	1	1	0.4





Problema 2. Organizar los horarios de partidos de La Liga(IV)

 Es posible la coincidencia de horarios pero en este caso la audiencia de cada partido se verá afectada y se estima que se reduce en porcentaje según la siguiente tabla dependiendo del número de coincidencias:

Coincidencias	-%
0	0%
1	25%
2	45%
3	60%
4	70%
5	75%
6	78%
7	80%
8	80%





Problema 2. Organizar los horarios de partidos de La Liga(IV)

Los cálculos asociados a una jornada de ejemplo se realizan según se muestra en la siguiente tabla:

114	Pa	rtido	Categorías	Horario	Base(Mill.)	Ponderación	Base*Ponderación	Correción Coincidencia
Celta	-	Real Madrid	B-A	V20	1,3	0,4	0,52	0,52
Valencia	-	R. Sociedad	B-A	S12	1,3	0,55	0,72	0,72
Mallorca	-	Eibar	C-C	S16	0,47	0,7	0,33	0,33
Athletic	-	Barcelona	B-A	S18	1,3	0,8	1,04	1,04
Leganés	-	Osasuna	C-C	S20	0,47	1	0,47	0,47
Villarreal	1-1	Granada	B-C	D16	0,75	0,75	0,56	0,42
Alavés	-	Levante	B-B	D16	0,9	0,75	0,68	0,51
Espanyol	-	Sevilla	B-B	D18	0,9	0,85	0,77	0,77
Betis		Valladolid	B-C	D20	0,75	1	0-3	0,75
Atlético	-	Getafe	B-B	L20	0,9	0,4	0,36	0,36

Total: 5,88

 $=0,56 \times 0,75$

 $=0,68 \times 0,75$



Viu Universidad Internacional de Valencia

Problema 3. Combinar cifras y operaciones

- El problema consiste en analizar el siguiente problema y diseñar un algoritmo que lo resuelva.
- Disponemos de las 9 cifras del 1 al 9 (excluimos el cero) y de los 4 signos básicos de las operaciones fundamentales: suma(+), resta(-), multiplicación(*) y división(/)
- Debemos **combinarlos alternativamente sin repetir ninguno de ellos** para obtener una cantidad dada. Un ejemplo sería para obtener el 4:

$$4+2-6/3*1=4$$



Problema 3. Combinar cifras y operaciones

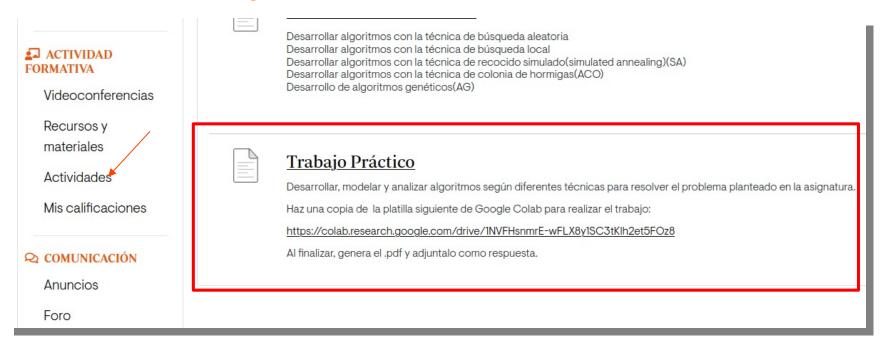
- Debe analizarse el problema para encontrar todos los valores enteros posibles planteando las siguientes cuestiones:
 - ¿Qué valor **máximo y mínimo** se pueden obtener según las condiciones del problema?
 - ¿Es posible encontrar todos los valores enteros posibles entre dicho mínimo y máximo ?
- Nota: Es posible usar la función de python "eval" para evaluar una expresión:

```
expresion = "4-2+6/3*1"
print(eval(expresion))
4.0
```





Problema del Trabajo Práctico



Plantilla: https://colab.research.google.com/drive/1NVFHsnmrE-wFLX8y1SC3tKlh2et5FOz8



Problema del Trabajo Práctico

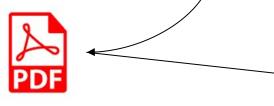
- Importante!. Además de resolver el problema será necesario responder algunas preguntas relacionadas con:
 - Complejidad
 - Justificar la estructura de datos elegida
 - Generar y probar con diferentes juegos de datos de entrada
- Fecha limite de entrega 1ª convocatoria: 19/07/2024
- Fecha limite de entrega 2ª convocatoria: 27/09/2024

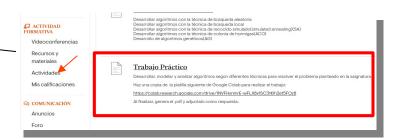
Generar un Notebook en GitHub (carpeta SEMINARIO)





Entrega de documento .pdf con en Notebook (como las A. Guiadas)







Plantilla para el documento

https://colab.research.google.com/drive/1NVFHsnmrE-wFLX8y1SC3tKlh2et5FOz8



Branch: master	03MAIRAlgoritmos-de-Optimizacion2019 / SEMINARIO / Seminario(plantilla).ipynb					Find file	Сору	path
🛖 raul27868	Creado con Colaboratory				3	3dd96ea 4	minutes	s ago
1 contributor								
417 lines (41	.7 sloc) 8.96 KB	<>		Raw	Blame	History		Î
	CO Open in Colab							
	Algoritmos de optimización - Seminario							
	Nombre y Apellidos: Url: https://aithub.com//03MAIRAlgoritmos-de-Optimizacion2019/tree/master/SEM	INARIO	2					
	Problema:	11474141	2					
	Elección de grupos de población homogéneos							
	Organizar los horarios de partidos de La Liga							
	3. Combinar cifras y operaciones							
	Descripción del problema:(copiar enunciado)							
	•••							
	(*) La respuesta es obligatoria							
In [0]:								
	(*)¿Cuantas posibilidades hay sin tener en cuenta las restricciones?							
	¿Cuantas posibilidades hay teniendo en cuenta todas las restricciones.							





Cabecera



- Algoritmos de optimización - Seminario

Nombre y Apellidos:	
Url https://github.com//03MAIRAlgoritmos-de-Optimizacion	-2019/tree/master/SEMINARIO
Problema:	
 Elección de grupos de pobleción homogéneos Organizar los horarios de partidos de La Liga Combinar cifras y operaciones 	Añadir texto del enunciado
Descripción del problema:(copiar enunciado)	Allauli texto del elluliciado
(*) La respuesta es obligatoria	





Pregunta – Respuesta (texto + Python)



(*)¿Cuantas posibilidades hay sin tener en cuenta las restricciones?					
¿Cuantas posibilidades hay teniendo en cuenta todas las restricciones.					
Respuesta					
Texto	Obligatoria				
[] Código python					





Pregunta – Respuesta (texto + Python)



(*)¿Cuantas posibilidades hay sin tener en cuenta las restricciones?

¿Cuantas posibilidades hay teniendo en cuenta todas las restricciones.

Respuesta:

Para calcular el numero posible de soluciones es necesario algo de combinatoria y saber contar.

Suponemos que el número de ciudades es N y que partimos de una ciudad dada. Podemos suponer un procedimiento que vaya construyendo todas las soluciones.

Para el primer viaje disponemos de N-1 ciudades ya que debemos eliminar la ciudad de partida como posible candidata. Para la segunda ciudad a visitar disponemos de N-2 posiblidades. Por tanto ya tenemos (N-1)x(N-2) para visitar 2 ciudades.

Si seguimos el razonamiento deducimos que hay (N-1)! (factoria de N-1) posibilidades.

Puesto que el camino es circular(comienza y termina en la misma ciudad) debemos tener en cuenta que cada ruta tiene una ruta inversa semejante. La primera se convierte en la última, la segunda en la penultima y así sucesivamente. Por tanto sin no queremos tener en cuenta esta repeticion en total tenemos (N-1)!/2





Problema del Trabajo Práctico. Preguntas(1/3)

Pregunta – Respuesta (texto + Python)



- (*)¿Cuantas posibilidades hay sin tener en cuenta las restricciones?
- ¿Cuantas posibilidades hay teniendo en cuenta todas las restricciones.
- (*) ¿Cual es la estructura de datos que mejor se adapta al problema? Argumenta la respuesta (Es posible que hayas elegido una al principio y veas la necesidad de cambiar, argumenta)
- (*)¿Cual es la función objetivo?
- (*)¿Es un problema de maximización o minimización?



(*) Obligatorio responder



Problema del Trabajo Práctico. Preguntas(2/3)

Pregunta – Respuesta (texto + Python)



- Diseña un algoritmo para resolver el problema por fuerza bruta
- Calcula la complejidad del algoritmo por fuerza bruta
- (*)Diseña un algoritmo que mejore la complejidad del algoritmo por fuerza bruta. Argumenta porque crees que mejora el algoritmo por fuerza bruta
- (*)Calcula la complejidad del algoritmo
- Según el problema (y tenga sentido), diseña un juego de datos de entrada aleatorio.





Problema del Trabajo Práctico. Preguntas (3/3)

Pregunta – Respuesta (texto + Python)

30%

- Aplica el algoritmo al juego de datos aleatorio generado.
- Enumera las referencias que has utilizado(si ha sido necesario) para llevar a cabo el trabajo
- Describe brevemente en unas líneas como crees que es posible avanzar en el estudio del problema. Ten en cuenta incluso posibles variaciones del problema y/o variaciones al alza del tamaño.





Problema del Trabajo Práctico. Evaluación.

Total 13 cuestiones:

6 obligatorias(*), aseguran 7/10

- 7 opcionales, añaden 2 puntos más: 9/10
- 1 punto por presentación, descripción, ...
 - lenguaje claro
 - código comentado
 - acompaña ilustraciones(imágenes) si es necesario

. . .

Fecha limite de entrega 1ª convocatoria: 19/07/2024

Fecha limite de entrega 2ª convocatoria: 27/09/2024



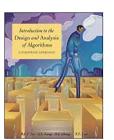


Ampliación de conocimientos y habilidades

- Bibliografía
 - -Brassard, G., y Bratley, P. (1997). Fundamentos de algoritmia. ISBN 13: 9788489660007
 - -Guerequeta, R., y Vallecillo, A. (2000). Técnicas de diseño de algoritmos. http://www.lcc.uma.es/~av/Libro/indice.html
 - -Lee, R. C. T., Tseng, S. S., Chang, R. C., y Tsai, Y. T. (2005). Introducción al diseño y análisis de algoritmos. *ISBN* 13: 9789701061244
- Pérez Aguila, R. (2012). Una introducción a las matemáticas para el análisis y diseño de algoritmos. ISBN 13: 9781413576474. https://tinyurl.com/yzlt5oed





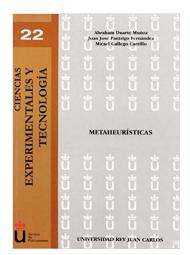


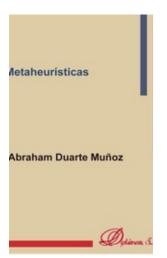


Ampliación de conocimientos y habilidades

- Bibliografía
 - Duarte, A. (2008). Metaheurísticas. Madrid: Dykinson.

https://elibro-net.universidadviu.idm.oclc.org/es/ereader/universidadviu/35696







Próximo día, prácticas sobre:

- Programación dinámica
- Búsqueda en grafos, ramificación y poda.
- Descenso del gradiente



 $Pg.: \langle N^o \rangle$

¿Preguntas?





 $Pg.: \langle N^o \rangle$

Gracias

jcamacho@professor.universidadviu.com

