# Análisis de Algoritmos Lectura: Técnicas de diseño de algoritmos: Programación Dinámica

Rodolfo Conde rodolfo.conde@itam.mx

Instituto Tecnológico Autónomo de México Maestría en Ciencias de la Computación

16 de octubre de 2020





Técnica: Programación Dinámica

Ejemplo: Multiplicar cadenas de matrices

Ejemplo: Un problema de Grafos

Notas y ejercicios





# Diseño y arquitectura de algoritmos

En esta lectura continuamos practicando las técnicas de análisis de algoritmos en general y aprenderemos una nueva técnica de diseño de algoritmos: La Programación Dinámica. Esta técnica es muy parecida a la de Dividir y Conquistar, al dividir el problema en subproblemas y ser de naturaleza recursiva. Sin embargo, los subproblemas en la Programación Dinámica tienden a tener subsubproblemas en común, lo cual da lugar para aplicar otras técnicas y herramientas para optimizar los algoritmos. Nuestra referencia principal es [Cor+09].





Técnica: Programación Dinámica

Ejemplo: Multiplicar cadenas de matrices

Ejemplo: Un problema de Grafos

Notas y ejercicios





# Arquitectura y diseño de algoritmos

Como mencionamos en la presentación pasada, en general nadie nos obliga a seguir un método especifico para construir un algoritmo, hay metodologías que nos ayudan a encontrar (o aproximarnos, en su defecto) al algoritmo más eficiente posible. Una de ellas es la Programación Dinámica. De hecho, es muy posible que en una forma indirecta, ya hayamos usado esta técnica para resolver algún problema.





P Dinámica





La Programación Dinámica es una técnica de diseño de algoritmos que resuelve los problemas al dividirlos en subproblemas (como en DyC), pero

Los subproblemas no son ajenos, comparte subsubproblemas.





La Programación Dinámica es una técnica de diseño de algoritmos que resuelve los problemas al dividirlos en subproblemas (como en DyC), pero

- Los subproblemas no son ajenos, comparte subsubproblemas.
- Para optimizar el algoritmo, se calculan las soluciones de los subsubproblemas y se guardan en una tabla.





La Programación Dinámica es una técnica de diseño de algoritmos que resuelve los problemas al dividirlos en subproblemas (como en DyC), pero

- Los subproblemas no son ajenos, comparte subsubproblemas.
- Para optimizar el algoritmo, se calculan las soluciones de los subsubproblemas y se guardan en una tabla.
- La tabla nos ahorra tiempo al no volver a recalcular la solución de un subsubproblema.





La Programación Dinámica es una técnica de diseño de algoritmos que resuelve los problemas al dividirlos en subproblemas (como en DyC), pero

- Los subproblemas no son ajenos, comparte subsubproblemas.
- Para optimizar el algoritmo, se calculan las soluciones de los subsubproblemas y se guardan en una tabla.
- La tabla nos ahorra tiempo al no volver a recalcular la solución de un subsubproblema.
- Intercambio Tiempo-Memoria.





## Programación Dinámica: Tipos de problemas

La Programación Dinámica se aplica de forma típica a Problemas de Optimización. Estos problemas

- Pueden tener muchas soluciones.
- Se busca la solución más óptima (máximo ó mínimo).
- Puede haber más de una solución óptima.









Al desarrollar un algoritmo por el método de Programación Dinámica, es muy común seguir los siguientes pasos:

1. Caracterizar la estructura de una solución óptima.





- 1. Caracterizar la estructura de una solución óptima.
- 2. Recursivamente definir el valor de una solución óptima.





- 1. Caracterizar la estructura de una solución óptima.
- 2. Recursivamente definir el valor de una solución óptima.
- 3. Calcular el valor (costo) de una solución óptima.





- 1. Caracterizar la estructura de una solución óptima.
- 2. Recursivamente definir el valor de una solución óptima.
- 3. Calcular el valor (costo) de una solución óptima.
- 4. Construir una solución óptima a partir de la información calculada.





Ahora veamos unos ejemplos para ver en acción a la Programación Dinámica.





00000

#### A continuación

Técnica: Programación Dinámica

Ejemplo: Multiplicar cadenas de matrices

Ejemplo: Un problema de Grafos

Notas y ejercicios





## Multiplicación de matrices

Como sabemos, multiplicar matrices es una operación fundamental para resolver diversos problemas con algoritmos.





Usualmente nos enfrentamos al problema de multiplicar dos matrices Cuadradas  $A_1, A_2 \in M_{n \times n}(\mathbb{R})$ 

$$A_1 \cdot A_2$$



## Multiplicación de matrices

Pero a veces también necesitamos calcular la multiplicación de una cantidad arbitraria de matrices  $A_1, A_2, \ldots, A_q$  tal que  $A_i$  tiene dimensiones  $n_i \times m_i$  and  $m_i = n_{i+1}$  para  $1 \le i < q$ .

$$A_1 \cdot A_2 \cdot \cdot \cdot A_q$$





La multiplicación de matrices es asociativa, entonces en si, no importa en que orden efectuemos la multiplicación, siempre obtendremos el mismo resultado.

$$A_1 \cdot A_2 \cdot \cdot \cdot A_q$$





## Multiplicación de matrices

Pero. . .

$$A_1 \cdot A_2 \cdot \cdot \cdot A_q$$





Pero. . .¿Cambia la cantidad de multiplicaciones escalares que debemos realizar?

$$A_1 \cdot A_2 \cdot \cdot \cdot A_q$$





## Ejemplo 1

#### Consideremos:

$$A_1 \in M_{100 \times 20}(\mathbb{R}),$$

$$A_2 \in M_{20 \times 20}(\mathbb{R}),$$

$$A_3 \in M_{20 \times 200}(\mathbb{R}).$$





## Multiplicación de matrices

#### Ejemplo 1

Consideremos:

$$A_1 \in M_{100 \times 20}(\mathbb{R}),$$

$$A_2 \in M_{20 \times 20}(\mathbb{R}),$$

$$A_3 \in M_{20 \times 200}(\mathbb{R}).$$

Y las posibles formas de asociar la multiplicación:

$$(A_1(A_2A_3))$$
 y  $((A_1A_2)A_3)$ 





#### Ejemplo 1

#### Consideremos:

$$A_1 \in M_{100 \times 20}(\mathbb{R}),$$

$$A_2 \in M_{20 \times 20}(\mathbb{R}),$$

$$A_3 \in M_{20 \times 200}(\mathbb{R}).$$

Para la primera opción:

$$(A_1(A_2A_3))$$





## Multiplicación de matrices

#### Ejemplo 1

Consideremos:

$$A_1 \in M_{100 \times 20}(\mathbb{R})$$
,

$$A_2 \in M_{20 \times 20}(\mathbb{R})$$
,

$$A_3 \in M_{20 \times 200}(\mathbb{R}).$$

Para la primera opción:

$$(A_1 \underbrace{(A_2A_3)}_{20\cdot 20\cdot 200=80000})$$





#### Ejemplo 1

Consideremos:

$$A_1 \in M_{100 \times 20}(\mathbb{R})$$
,

$$A_2 \in M_{20 \times 20}(\mathbb{R}),$$

$$A_3 \in M_{20 \times 200}(\mathbb{R}).$$

Para la primera opción:

$$\underbrace{(A_1(A_2A_3))}_{100\cdot 20\cdot 200=400000}$$





# Multiplicación de matrices

#### Ejemplo 1

#### Consideremos:

$$A_1 \in M_{100 \times 20}(\mathbb{R}),$$

$$A_2 \in M_{20 \times 20}(\mathbb{R}),$$

$$A_3 \in M_{20 \times 200}(\mathbb{R}).$$

Para la primera opción: El total de operaciones es  $80000 \cdot 400000 = 32000000000$ 

$$(A_1(A_2A_3))$$





#### Ejemplo 1

#### Consideremos:

$$A_1 \in M_{100 \times 20}(\mathbb{R}),$$

$$A_2 \in M_{20 \times 20}(\mathbb{R})$$
.

$$A_3 \in M_{20 \times 200}(\mathbb{R}).$$

Un calculo similar para la segunda opción nos da un total de operaciones de  $40000 \cdot 400000 = 16000000000$ 

$$((A_1A_2)A_3)$$





# Ejemplo 1

#### Consideremos:

$$A_1 \in M_{100 \times 20}(\mathbb{R})$$
,

$$A_2 \in M_{20 \times 20}(\mathbb{R})$$
,

$$A_3 \in M_{20 \times 200}(\mathbb{R}).$$

¡¡La segunda opción es mejor!!

$$((A_1A_2)A_3)$$





El problema de Multiplicar matrices en cadena consta de los siguientes datos:





El problema de Multiplicar matrices en cadena consta de los siguientes datos:

Entrada: Una secuencia de enteros  $\langle p_0, \dots, p_n \rangle$  que representan la secuencia de matrices  $\langle A_1, \dots, A_n \rangle$  tal que  $A_i \in M_{p_{i-1} \times p_i}(\mathbb{R})$ .





El problema de Multiplicar matrices en cadena consta de los siguientes datos:

- Entrada: Una secuencia de enteros  $\langle p_0, \ldots, p_n \rangle$  que representan la secuencia de matrices  $\langle A_1, \dots, A_n \rangle$  tal que  $A_i \in M_{p_{i-1} \times p_i}(\mathbb{R})$ .
- Salida: La forma de asociar el producto de todas las matrices, de tal manera que se minimice la cantidad de multiplicaciones escalares necesarias para calcular el producto de todas las matrices.





El problema de Multiplicar matrices en cadena consta de los siguientes datos:

Entrada: Una secuencia de enteros  $\langle p_0, \dots, p_n \rangle$  que representan la secuencia de matrices  $\langle A_1, \dots, A_n \rangle$  tal que  $A_i \in M_{p_{i-1} \times p_i}(\mathbb{R})$ .

Salida: La forma de asociar el producto de todas las matrices, de tal manera que se minimice la cantidad de multiplicaciones escalares necesarias para calcular el producto de todas las matrices.

Nota: No se pide multiplicar las matrices.





La primera opción es verificar por busqueda exaustiva. ¿De cuantas formas se pueden poner los parantesis para una secuencia de n matrices? Sea P(n) el número de formas posibles.





Si 
$$n = 1$$
, solo hay una forma.  $P(1) = 1$ 



### Multiplicación de matrices: Primer análisis

n > 1

Cuando hemos seleccionado donde poner los primeros parentesis, encerrando k matrices  $A_i \cdots A_{i+k}$ , debemos escoger una forma de poner parentesis para las otras n-k matrices.





#### n > 1

Cuando hemos seleccionado donde poner los primeros parentesis, encerrando k matrices  $A_i \cdots A_{i+k}$ , debemos escoger una forma de poner parentesis para las otras n-k matrices. Y esto lo hacemos para todo  $k=1,\ldots,n-1$ .





$$P(n) = \begin{cases} 1 & n = 1\\ \sum_{k=1}^{n-1} P(k)P(n-k) & n > 1 \end{cases}$$
 (1)





$$P(n) = \begin{cases} 1 & n = 1\\ \sum_{k=1}^{n-1} P(k)P(n-k) & n > 1 \end{cases}$$
Una recurrencia. (1)



### Multiplicación de matrices: De cuantas formas

No es difícil ver que una solución a la recurrencia

$$P(n) = \begin{cases} 1 & n = 1\\ \sum_{k=1}^{n-1} P(k)P(n-k) & n > 1 \end{cases}$$
 (2)





### Multiplicación de matrices: De cuantas formas

No es difícil ver que una solución a la recurrencia

$$P(n) = \begin{cases} 1 & n = 1\\ \sum_{k=1}^{n-1} P(k)P(n-k) & n > 1 \end{cases}$$
 (2)

es (en pizarron)  $P(n) = \Omega(2^n)$ . (¡¡Exponencial!!)





¿DyC?



¿DyC? Tampoco es buena opción, pues estariamos recorriendo todo el arbol de formas posibles de asociar las matrices, lo cual esta relacionado con P(n).





Vamos a intentar resolverlo con Programación Dinámica (PD).





El primer paso es caracterizar la estructura de una solución óptima. Es decir, si tenemos una solución optima para multiplicar  $A_{i\cdots j}$  a partir de poner parentesis entre  $A_k$  y  $A_{k+1}$  para k tal que  $i \leq k < j$ 





El primer paso es caracterizar la estructura de una solución óptima. Es decir, si tenemos una solución optima para multiplicar  $A_{i\cdots j}$  a partir de poner parentesis entre  $A_k$  y  $A_{k+1}$  para k tal que  $i \leq k < j$  ¿Las formas en las que hemos colocado parentesis para multiplicar  $A_{i\cdots k}$  y  $A_{k+1\cdots j}$  son óptimas?





El primer paso es caracterizar la estructura de una solución óptima. Es decir, si tenemos una solución optima para multiplicar  $A_{i\cdots j}$  a partir de poner parentesis entre  $A_k$  y  $A_{k+1}$  para k tal que  $i \leq k < j$  ¿Las formas en las que hemos colocado parentesis para multiplicar  $A_{i\cdots k}$  y  $A_{k+1\cdots j}$  son óptimas? (Vamonos al pizarrón).









$$m[i,j] = \begin{cases} 0 & i = j, \\ m[i,k] + m[k+1,j] + p_{i-1}p_kp_j & i < j. \end{cases}$$
(3)





$$\underbrace{m[i,j]}_{\text{Óptimo de }A_{i\cdots i}} = \begin{cases}
0 & i = j, \\
m[i,k] + m[k+1,j] + p_{i-1}p_kp_j & i < j.
\end{cases} (3)$$





$$m[i,j] = \begin{cases} 0 & i = j, \\ \underbrace{m[i,k]}_{\text{Óptimo de } A_{i\cdots k}} + m[k+1,j] + p_{i-1}p_kp_j & i < j. \end{cases}$$
(3)





#### PD Paso 2

$$m[i,j] = \begin{cases} 0 & i = j, \\ m[i,k] + \underbrace{m[k+1,j]}_{\text{Optimo de } A_{k+1\cdots j}} + p_{i-1}p_kp_j & i < j. \end{cases}$$
(3)





$$m[i,j] = \begin{cases} 0 & i = j, \\ m[i,k] + m[k+1,j] + \underbrace{p_{i-1}p_kp_j}_{\text{Costo de } A_{i\cdots k} \cdot A_{k+1\cdots j}} & i < j. \end{cases}$$
(3)





### PD Paso 2

El siguiente paso es encontrar una ecuación recursiva que define el valor de una solución óptima. Esta ecuación será:

$$m[i,j] = \begin{cases} 0 & i = j, \\ m[i,k] + m[k+1,j] + p_{i-1}p_kp_j & i < j. \end{cases}$$
(3)

¡¡ Suponiendo que conocemos k!!





$$m[i,j] = \begin{cases} 0 & i = j, \\ \min_{i \le k < j} \{ m[i,k] + m[k+1,j] + p_{i-1}p_kp_j \} & i < j. \end{cases}$$
(3)

¿Cómo encontrar k para cada subproblema?





$$m[i,j] = \begin{cases} 0 & i = j, \\ \min_{i \le k < j} \{ m[i,k] + m[k+1,j] + p_{i-1}p_k p_j \} & i < j. \end{cases}$$
(3)

- ¿Cómo encontrar k para cada subproblema?
- ¿Cuántas posibilidades hay en total?





$$m[i,j] = \begin{cases} 0 & i = j, \\ \min_{i \le k < j} \{ m[i,k] + m[k+1,j] + p_{i-1}p_k p_j \} & i < j. \end{cases}$$
(3)

- ¿Cómo encontrar k para cada subproblema?
- ¿Cuántas posibilidades hay en total?





#### PD Paso 3

El siguiente paso es calcular el costo de una solución óptima.





El siguiente paso es calcular el costo de una solución óptima.

► Solución recursiva → Exponencial





#### PD Paso 3

El siguiente paso es calcular el costo de una solución óptima.

- ▶ Solución recursiva → Exponencial
- Recordemos soluciones calculadas





### PD Paso 3

El siguiente paso es calcular el costo de una solución óptima.

- ▶ Solución recursiva → Exponencial
- Recordemos soluciones calculadas
- Estrategía de abajo hacia arriba<sup>1</sup>.



<sup>&</sup>lt;sup>1</sup>bottom-up

# PD Paso 3: Algoritmo I

```
function Matrix-Chain-Order(p) //p = \langle p_0, \dots p_n \rangle
       n = p.lenght - 1
        Let m[1 \cdots n, 1 \cdots n] be a new table
3:
        for i = 1 to n do
4:
            m[i,i] = 0 // Costo de la cadena A_i
 5:
        end for
6:
7:
        for l=2 to n do
8:
            for i = 1 to n - l + 1 do
               i = i + l - 1
9:
               m[i,j] = \infty
10:
               for k = i to i - 1 do
11:
                   q = m[i, k] + m[k + 1, j] + p_{i-1}p_kp_i
12:
```





```
      13:
      if q < m[i,j] then

      14:
      m[i,j] = q

      15:
      end if

      16:
      end for

      17:
      end for

      18:
      end for

      19:
      return m

      20:
      end function
```





```
function M-C-O(p)
   n = p.lenght - 1
   Let m[1 \cdots n, 1 \cdots n] be a new table
   for i = 1 to n do
       m[i, i] = 0
   end for
   for l=2 to n do
       for i = 1 to n - l + 1 do
          i = i + l - 1
          m[i,j] = \infty
          for k = i to i - 1 do
                 = m[i,k] + m[k+1,j] +
p_{i-1}p_kp_i
              if q < m[i,j] then
                 m[i,j]=q
              end if
          end for
       end for
   end for
   return m
end function
```

Vamos a analizar como funciona este código (Pizarrón)



```
function M-C-O(p)
   n = p.lenght - 1
   Let m[1 \cdots n, 1 \cdots n] be a new table
   for i = 1 to n do
       m[i, i] = 0
   end for
   for l=2 to n do
       for i = 1 to n - l + 1 do
          i = i + l - 1
          m[i,j] = \infty
                                                  ¿Cual es la complejidad de
          for k = i to i - 1 do
                                                    tiempo del algoritmo?
                = m[i,k] + m[k+1,j] +
p_{i-1}p_kp_i
              if q < m[i,j] then
                 m[i,j]=q
              end if
          end for
       end for
   end for
   return m
```



end function

4 □ > 4 □ > 4 □ > 4 □ >

```
function M-C-O(p)
   n = p.lenght - 1
   Let m[1 \cdots n, 1 \cdots n] be a new table
   for i = 1 to n do
       m[i, i] = 0
   end for
   for l=2 to n do
       for i = 1 to n - l + 1 do
          i = i + l - 1
          m[i,j] = \infty
                                                               O(n^3).
          for k = i to i - 1 do
                 = m[i,k] + m[k+1,j] +
p_{i-1}p_kp_i
              if q < m[i,j] then
                  m[i,j] = q
              end if
          end for
       end for
   end for
   return m
end function
```



```
function M-C-O(p)
   n = p.lenght - 1
   Let m[1 \cdots n, 1 \cdots n] be a new table
   for i = 1 to n do
       m[i, i] = 0
   end for
   for l=2 to n do
       for i = 1 to n - l + 1 do
          i = i + l - 1
                                                    Se puede probar que de
          m[i,j] = \infty
                                                  hecho el algoritmo corre en
          for k = i to i - 1 do
                = m[i,k] + m[k+1,j] +
                                                         tiempo \Omega(n^3)
p_{i-1}p_kp_i
              if q < m[i,j] then
                 m[i,j]=q
              end if
          end for
       end for
   end for
   return m
```



end function

4 □ > 4 □ > 4 □ > 4 □ >

```
function M-C-O(p)
   n = p.lenght - 1
   Let m[1 \cdots n, 1 \cdots n] be a new table
   for i = 1 to n do
       m[i, i] = 0
   end for
   for l=2 to n do
       for i = 1 to n - l + 1 do
          i = i + l - 1
          m[i,j] = \infty
                                                  ¿Cual es la complejidad de
          for k = i to i - 1 do
                                                    espacio del algoritmo?
                = m[i,k] + m[k+1,j] +
p_{i-1}p_kp_i
              if q < m[i,j] then
                 m[i,j]=q
              end if
          end for
       end for
   end for
   return m
```





end function

```
function M-C-O(p)
   n = p.lenght - 1
   Let m[1 \cdots n, 1 \cdots n] be a new table
   for i = 1 to n do
       m[i, i] = 0
   end for
   for l=2 to n do
       for i = 1 to n - l + 1 do
          i = i + l - 1
           m[i,j] = \infty
                                                                \Theta(n^2)
           for k = i to i - 1 do
                 = m[i,k] + m[k+1,j] +
p_{i-1}p_kp_i
              if q < m[i,j] then
                  m[i,j] = q
              end if
           end for
       end for
   end for
   return m
end function
```



El último paso en nuestra solución con PD es calcular una solución optima que nos da el costo óptimo calculado en el paso 3. En el caso del problema de Multiplicar matrices en cadena, la solución óptima es una forma de acomodar los parentesis para efectuar la multiplicación de la cadena de matrices, de tal forma que el número de operaciones escalares sea de costo óptimo.





#### A continuación

Técnica: Programación Dinámica

Ejemplo: Multiplicar cadenas de matrices

Ejemplo: Un problema de Grafos

Notas y ejercicios





El camino más largo simple dirigido entre un par de nodos en un grafo dirigido acíclico.





## Un problema común en Grafos

Entrada: Un grafo dirigido acíclico D=(N,A) con pesos en los arcos  $(p\colon A\to \mathbb{R})$  y dos nodos distinguidos  $s,t\in N$ .

Salida: Un camino simple de *s* a *t* que maximice la suma de pesos de los arcos.





Primero debemos mostrar que las instancias optimas de este problema muestran subestructura optima (Pizarrón)





El siguiente paso es construir una ecuación recursiva para calcular el costo de una solución óptima (Pizarrón).





Ahora debemos construir un algoritmo para calcular el costo óptimo de una solución óptima.





```
Require: D = (N, A) un grafo dirigido acíclico
1: function Longest-Simple-Path(D, p, s, t)
2:
        q = -\infty
3:
        for all v \in \text{outNeigbors}(s) do
4:
            w_1 = p(\overrightarrow{sv}) + \text{Longest-Simple-Path}(D, p, v, t)
5:
            if q < w_1 then
6:
                q = w_1
7:
            end if
8:
        end for
9:
        return q
    end function
```





# Camino simple más largo: PD Paso 3

```
Require: D = (N, A) un grafo dirigido acíclico
      // De arriba hacia abajo con memoización (top-down, memoization)
 1: function Longest-Simple-Path-PD(D, p, s, t)
2:
        n = |N|
3:
        Let w[1 \cdots n, 1 \cdots n] be a new table
4:
        for all i, j \in N do
            if \overrightarrow{ii} \in A then
5:
                w[i,j] = p(\overrightarrow{ij})
6:
7:
            else if i == j then
8:
                w[i, j] = 0
9:
            else
10:
                w[i,j] = -\infty
11:
            end if
12:
         end for
13:
         return Longest-Simple-Path-PD-AUX(D, p, s, t, w)
14: end function
```



```
function Longest-Simple-Path-PD-AUX(D, p, s, t, w)
2:
       if w[s,t] \ge 0 then
 3:
           return w[s, t]
4:
       else
5:
           q = -\infty
6:
           for all v \in \text{outNeigbors}(s) do
7:
               w_1 = w[s, v] + \text{Longest-Simple-Path-PD-AUX}(D, p, v, t, w)
8:
               if q < w_1 then
9:
                   a = w_1
10:
               end if
11:
            end for
12:
            w[s,t]=q
13:
        end if
14:
        return q
15: end function
```



Técnica: Programación Dinámica

Ejemplo: Multiplicar cadenas de matrices

Ejemplo: Un problema de Grafos

Notas y ejercicios





4 D > 4 A > 4 B > 4 B >

## Ejercicios del capítulo I

#### **EJERCICIO 1**

Construye en pseudocódigo un algoritmo para multiplicar dos matrices de dimensiones arbitrarias y compatibles. Tu programa debe estar preparado para el caso en que las matrices no sean compatibles y mandar un mensaje de error. ¿Cuál es la complejidad de tiempo de tu algoritmo?

#### **EJERCICIO 2**

Considera una variante del problema de Multiplicar matrices en cadena, en la cual en lugar de minimizar la cantidad de multiplicaciones escalares, se busca maximizar esta cantidad. ¿Este problema exibe subestructura óptima? Justifica.





Thomas H. Cormen y col. *Introduction to Algorithms, Third Edition*. 3rd. The MIT Press, 2009. ISBN: 0262033844, 9780262033848.



