Experimentación temporal sobre a complexidade algorítmica

Pablo Gil Pérez – pablo.gil.perez@rai.usc.es

Introdución:

Neste informe analizarase de maneira experimental o costo temporal da implementación de dous algoritmos de distinta complexidade coa finalidade de observar a súa relación co costo temporal da súa execución e como este coste evoluciona a medida que o fai a talla do problema.

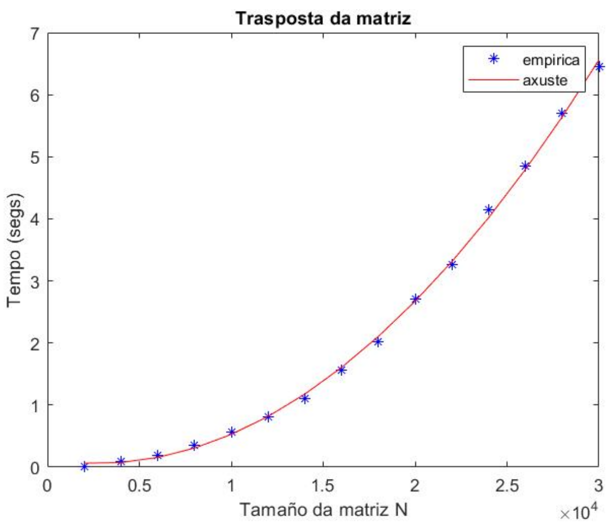
As implementacións a analizar están baseadas en dous algoritmos distintos. En primeiro lugar trátase o algoritmo de trasposición de matrices (no noso caso traspoñemos unha matriz (n,n/2)), o cal ten complexidade cuadrática, e en segundo lugar o algoritmo de multiplicación de matrices, (no noso caso multiplicamos unha matriz (n,n) por outra (n,n/2)) que ten unha complexidade cúbica.

No primeiro caso contamos con complexidade cuadrática debido a que se teñen que recorrer todas as filas e todas as columnas dunha matriz (n,n/2), e como ambos parámetros dependen de n estanse recorrendo $n*(n/2)=n^2/2$ elementos, valor que como podemos observar, crece cuadráticamente. No caso da multiplicación temos que ter en conta que ademais de recorrer filas e columnas, por cada elemento recórrese á súa vez unha fila do primeiro operando e unha columna do segundo, polo que estamos recorrendo n filas da primeira matriz por n/2 columnas da segunda por n columnas da primeira e filas da segunda, en total $n*n/2*n=n^3/2$, valor que como podemos observar crece de maneira cúbica.

Para comezar a investigación, realizáronse as dúas implementacións dos algoritmos de xeito que a súa complexidade non superase a complexidade do propio algoritmo. Unha vez feita a implementación pasouse á recollida dos datos temporais, tendo en conta unicamente o tempo que transcorre dende que comezan as operacións antes citadas ata que finalizan, sen contar outras operacións secundarias, para unha análise máis precisa. Estes datos correspondese cunha máquina en concreto, pero mediante a evolución dos propios valores podemos observar como crece o coste temporal do algoritmo. Unha vez feita a recollida dos datos pasouse á súa representación e análise. A gráfica fíxose mediante a ferramenta MATLAB®, representáronse os datos obtidos experimentalmente, e fíxose un axuste a través dun polinomio de grao igual á complexidade do algoritmo para ver de xeito máis preciso a evolución do coste. O seguinte paso é a análise:

Análise:

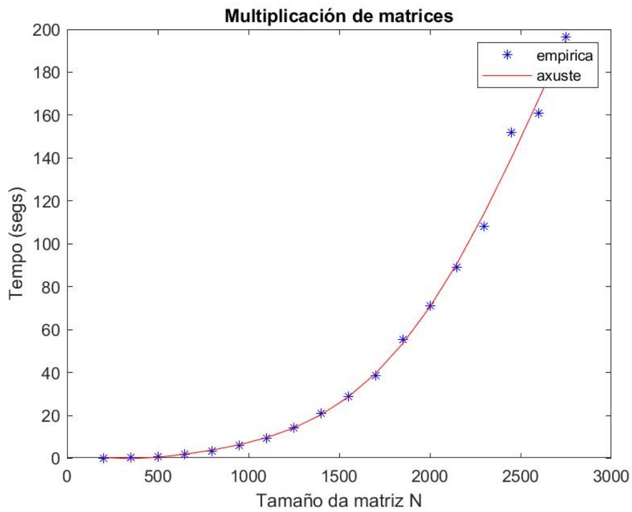
En primeiro lugar imos comezar coa análise do algoritmo da trasposición. Conta cunha complexidade cuadrática e a gráfica que se obtivo a partir dos datos experimentais é a seguinte:



Como podemos observar, o axuste obtido a través dos datos segue unha traxectoria parabólica crecente, chegando a rexistrarse unha tardanza de máis de 6 segundos no último caso. É impensable tratar con datos desta magnitude mediante este algoritmo debido ó alto coste temporal. Podemos definir o tamaño máximo tratable por este algoritmo no caso desta máquina como unha talla entre 20000 e 25000, é dicir, que non pase dos 3-4 segundos.

O primeiro valor que se colle na gráfica é 2000, e logo vaise avanzando de 2000 en 2000 ata chegar a 30000. Comézase en 2000 porque con valores máis baixos o tempo que se tarda en realizar a operación é moi reducido e non axudaría a apreciar a evolución. Por outro lado, chegamos ata 30000 porque a partir deste valor, xa comezamos a obter datos temporais elevados e pouco tratables. Por último, o paso escollido é de 2000 para que exista un equilibrio entre o número de puntos que representamos e non haxa demasiados, á vez que sexan suficientes para observar o crecemento.

Pasamos agora a analizar o caso da multiplicación. A complexidade deste algoritmo é cúbica e a súa gráfica é a seguinte:



Neste caso podemos ver como a curva crece máis rápido. Este diferencia nótase especialmente para tallas elevadas, e se nos fixamos nos valores temporais podemos observar como, en comparación co caso anterior, para tallas máis pequenas temos valores temporais moito mais elevados. Isto ocorre porque ó ser maior a complexidade, a medida que aumentamos a talla crece moito máis rápido que un algoritmo de menor complexidade. Como consecuencia, a talla que denominamos como tratable é moito menor que no caso anterior. Temos un dato de 3.47 segundos para unha talla n=800, polo que podemos dicir que a talla máxima tratable por este algoritmo no caso desta gráfica está entre 800 e 900, para non superar excesivamente os 3-4 segundos.

A talla inicial que se toma en conta para a gráfica é de 200, de novo porque tomar valores inferiores non suporía ningunha mellora na avaliación do crecemento temporal ó ser tempos moi pequenos. O tamaño máximo é de 2750, sendo o tempo recollido moitísimo maior que na gráfica anterior. Fíxose desta maneira para apreciar mellor a rapidez do aumento temporal a medida que aumenta a talla e como con valores a priori non moi grandes, tardase moito. O paso é de 150, escollido así para nin ter poucos puntos experimentais que fagan a análise menos precisa, como para non ter demasiados puntos que fagan a gráfica menos lexible.

Conclusión:

A conclusión que podemos extraer é que a complexidade dun algoritmo infúe de xeito moi notable no rendemento que nos vai proporcionar, e aínda máis se tratamos tallas grandes. Por este motivo temos que intentar sempre utilizar algoritmos eficientes e coa menor complexidade posible para empregar o mínimo gasto temporal e tamén o menor gasto de memoria que se poida, tendo en conta tamén que en moitos casos a memoria e o tempo están relacionados e ten que optarse por unha solución de compromiso entre eles para non prexudicar gravemente ningún.