2.3. Algoritmo de Euclides

En los enteros y en general en los dominios euclidianos el Algoritmo de Euclides es una herramienta para calcular el máximo común divisor de dos enteros a y b. Este se basa en el algoritmo de la división.

Definición 2.3 Sean a y b enteros no ambos iguales a cero. El máximo común divisor de a y b es el mayor entero que es divisor común de a y b. Lo notamos m.c.d.(a,b) o simplemente (a,b).

Si $x|a$ entonces $x|(-a)$, luego

$$\text{m.c.d.}(a,b) = \text{m.c.d.}(a,-b) = \text{m.c.d.}(-a,b) = \text{m.c.d.}(-a,-b).$$

```
In[21]:= GCD[2^20 + 20^2, 123456]

Out[21]= 16
```

Como aplicación del algoritmo de la división tenemos el siguiente resultado.

Teorema 2.7 (Igualdad de Bezout) Sean a y b enteros no ambos iguales a cero. El máximo común divisor de a y b es el menor entero positivo que puede escribirse como combinación lineal de a y b, es decir de la forma $ax+by$ con x,y enteros.

Vamos a necesitar el siguiente lema.

Lema 2.1 Sean a,b,q y r enteros tales que $a = bq + r$, entonces m.c.d.$(a,b) =$ m.c.d.(b,r).

Sean $a,b \in \mathbb{Z}$ con $a > b > 0$, por el Algoritmo de la División tenemos que

$$\begin{aligned}
\overbrace{a}^{r_0} &= \overbrace{b}^{r_1} q_1 + r_2 \\
b &= r_2 q_2 + r_3 \\
r_2 &= r_3 q_3 + r_4 \\
&\vdots \\
r_{n-2} &= r_{n-1} q_{n-1} + r_n \\
r_{n-1} &= r_n q_n + \cancelto{0}{r_{n+1}}
\end{aligned} \tag{2.1}$$

Por el Lema 2.1 se concluye que m.c.d.$(a,b) = r_n$. El entero n, correspondiente al número de pasos para calcular el máximo común divisor de a y b mediante el Algoritmo de Euclides, se denota por Euc$(a,b) = n$.

Ejemplo 2.17 Utilizando el algoritmo de Euclides podemos calcular el máximo común divisor de $a = 98765$ y $b = 1234$. En efecto tenemos los siguientes pasos:

$$\begin{aligned} 98765 &= 1234 \cdot 80 + 45 \\ 1234 &= 45 \cdot 27 + 19 \\ 45 &= 19 \cdot 2 + 7 \\ 19 &= 7 \cdot 2 + 5 \\ 7 &= 5 \cdot 1 + 2 \\ 5 &= 2 \cdot 2 + 1 \\ 2 &= 1 \cdot 2 + 0 \end{aligned}$$

Esto nos permite concluir que $(98765, 1234) = 1$, es decir son primos relativos. El algoritmo de Euclides también nos permite encontrar una combinación lineal del máximo común divisor. En efecto se van reemplazando cada uno de los residuos desde la penúltima ecuación.

$$\begin{aligned} 1 &= 5 - 2 \cdot 2 \\ &= 5 - 2(7 - 5 \cdot 1) = 3(5) - 2(7) \\ &= 3(19 - 7 \cdot 2) - 2(7) = 3(19) - 8(7) \\ &= 3(19) - 8(45 - 19 \cdot 2) = 19(19) - 8(45) \\ &= 19(1234 - 45 \cdot 27) - 8(45) = 19(1234) - 521(45) \\ &= 19(1234) - 521(98765 - 1234 \cdot 80) = 41699(1234) + (-521)98765 \end{aligned}$$

Por lo tanto, $(98765, 1234) = 1 = 41699(1234) + (-521)98765$.

```
In[22]:= GCD[98765, 1234]

Out[22]= 1

In[23]:= ExtendedGCD[98765, 1234]

Out[23]= {1, {-521, 41699}}
```

El procedimiento descrito en el ejemplo anterior se puede generalizar. Considere $0 < b < a$ enteros y supongamos que tenemos las ecuaciones dadas en (2.1). Entonces se tiene que

$$\begin{aligned} \text{m.c.d.}(a, b) = r_n &= r_{n-2} - q_{n-1}r_{n-1} = r_{n-2} - q_{n-1}(r_{n-3} - q_{n-2}r_{n-2}) \\ &= -q_{n-1}r_{n-3} + (1 + q_{n-1}q_{n-2}) = \cdots \end{aligned}$$

En general se puede demostrar que r_n es combinación lineal de r_{n-2} y r_{n-1}, pero también es combinación lineal de r_{n-3} y r_{n-2}, y así sucesivamente se verifica que r_n es combinación lineal de $r_0 = a$ y $r_1 = b$ (teorema de Bezout). Existe una forma alternativa para determinar la combinación lineal sin tener que memorizar todos los cocientes intermedios q_i. Este

procedimiento se conoce como el Algoritmo Extendido de Euclides.

Considere $0 < b < a$ enteros y supongamos que tenemos las ecuaciones dadas en (2.1). Definimos $r_0 = a, r_1 = b, x_0 = 1,\ x_1 = 0,\ y_0 = 0,\ y_1 = 1$ y las fórmulas de recurrencia

$$\begin{aligned} x_i &= x_{i-2} - x_{i-1}q_{i-1}, \\ y_i &= y_{i-2} - y_{i-1}q_{i-1}, \end{aligned}$$

para $i = 2, \ldots, n$. Bajo estas condiciones tenemos:

$$ax_i + by_i = r_i \tag{2.2}$$

para $i = 1, 2, \ldots, n$. En efecto, para $i = 0, 1$ se cumple

$$\begin{aligned} ax_0 + by_0 &= (a)(1) + b(0) = a = r_0, \\ ax_1 + by_1 &= (a)(0) + b(1) = b = r_1. \end{aligned}$$

Supongamos que para calcular r_i tenemos que $r_j = ax_j + by_j$ para todo $j = 0, 1, \ldots, i-1$. Del algoritmo de Euclides se tiene que $r_i = r_{i-2} - r_{i-1}q_{i-1}$. De la hipótesis de inducción y de las fórmulas de recurrencia tenemos que

$$\begin{aligned} r_i = r_{i-2} - r_{i-1}q_{i-1} &= (ax_{i-2} + by_{i-2}) - (ax_{i-1} + by_{i-1})q_{i-1} \\ &= a(x_{i-2} - x_{i-1}q_{i-1}) + b(y_{i-2} - y_{i-1}q_{i-1}) \\ &= ax_i + by_i. \end{aligned}$$

Por lo tanto se tiene para $j = i$.

Ejemplo 2.18 Encontremos (1001, 275) y escribámoslo como combinación lineal de ellos.

$$\begin{aligned} 1001 &= (275)(3) + 176 \\ 275 &= (176)(1) + 99 \\ 176 &= (99)(1) + 77 \\ 99 &= (77)(1) + 22 \\ 77 &= (22)(3) + 11 \\ 22 &= (11)(2) + 0. \end{aligned}$$

Usando las fórmulas de recurrencia tenemos la siguiente tabla,

i	q_{i-1}	x_i	y_i	$ax_i + by_i$
0	–	1	0	1001
1	–	0	1	275
2	3	1	-3	176
3	1	-1	4	99
4	1	2	-7	77
5	1	-3	11	22
6	3	11	-40	11
7	2	-25	91	0

por lo tanto, $(1001, 275) = 11 = (1001)(11) + (275)(-40)$.

Observe que el cociente q_{i-1} de la división de r_{i-2} entre r_{i-1} no se necesita para calcular los enteros r_i, x_i, y_i, esto evita tener que guardar todos los cocientes.

Por último vamos a analizar la complejidad del algoritmo de Euclides.

Teorema 2.8 Sean $a, b \in \mathbb{Z}$ tales que $a > b > 0$ con $n = \text{Euc}(a, b)$. Entonces $a \geq F_{n+2}$ y $b \geq F_{n+1}$.

Demostración. La demostración es por inducción sobre n. Para $n = 1 = \text{Euc}(a, b)$, esto significa que el algoritmo consiste únicamente de una división. Es decir que

$$a = bq + \cancelto{0}{r_2}.$$

Como $a > b > 0$, entonces el valor más pequeño posible de b es $1 = F_2$ y para a es $2 = F_3$.

Supongamos que la afirmación es cierta para todo $i < n$ y demostremos su veracidad para n. Por el Algoritmo de la División existen enteros q_1 y r_2 tales que $a = bq_1 + r_2$, con $0 \leq r_2 < b$. Además, $\text{Euc}(b, r_2) = n - 1$. Por hipótesis de inducción,

$$b \geq F_{n+1} \quad \text{y} \quad r_2 \geq F_n.$$

Por lo tanto, $a = bq_1 + r_2 \geq b + r_2 \geq F_{n+1} + F_n = F_{n+2}$. Quedando demostrado por el PIM. □

Definición 2.4 Sean $(u, v), (u', v') \in \mathbb{Z} \times \mathbb{Z}$. La pareja (u, v) es menor o igual que (en el orden lexicográfico) (u', v') si $u < u'$ o $u = u'$ y $v \leq v'$.

Teorema 2.9 (Lamé) Dado un entero positivo n, la pareja de enteros positivos (a, b) $(a > b > 0)$ más pequeña en el orden lexicográfico que satisface que $n = \text{Euc}(a, b)$ es $(a, b) = (F_{n+2}, F_{n+1})$.

Demostración. Del Teorema 2.8 se tiene que $a \geq F_{n+2}$ y $b \geq F_{n+1}$. Solo falta ver que $\text{Euc}(F_{n+2}, F_{n+1}) = n$. Esto se obtiene de la definición de la sucesión de Fibonacci:

$$\begin{aligned} F_{n+2} &= F_{n+1} + \overbrace{F_n}^{r_2} \\ F_{n+1} &= F_n + \overbrace{F_{n-1}}^{r_3} \\ &\vdots \\ F_4 &= F_3 + \overbrace{F_2}^{r_n} \\ F_3 &= 2F_2 + 0, \end{aligned}$$

luego $\text{Euc}(F_{n+2}, F_{n+1}) = n$. □

Recordemos la fórmula de Binet para los números de Fibonacci:

$$F_n = \frac{1}{\sqrt{5}} \left[\left(\underbrace{\frac{1+\sqrt{5}}{2}}_{\alpha} \right)^n - \left(\underbrace{\frac{1-\sqrt{5}}{2}}_{\beta} \right)^n \right], \quad n \geq 0.$$

Teorema 2.10 (Cota para $\text{Euc}(a,b)$) Sean $a, b \in \mathbb{Z}$ con $a > b > 0$. Entonces,

$$\text{Euc}(a,b) < c_1 \log a + c_2 - 2 + c_3 \frac{1}{a},$$

donde $c_1 = 1/\log\alpha \approx 2.08$, $c_2 = \log\sqrt{5}/\log\alpha \approx 1.67$, $c_3 = 1/(\sqrt{5}\log\alpha) \approx 0.93$. También se tiene que

$$\text{Euc}(a,b) < c_1 \log b + c_2 - 1 + c_3 \frac{1}{b}.$$

Demostración. Sea $n = \text{Euc}(a,b)$, entonces $a \geq F_{n+2}$. Así

$$a \geq \frac{\alpha^{n+2} - \beta^{n+2}}{\sqrt{5}} > \frac{\alpha^{n+2} - 1}{\sqrt{5}}.$$

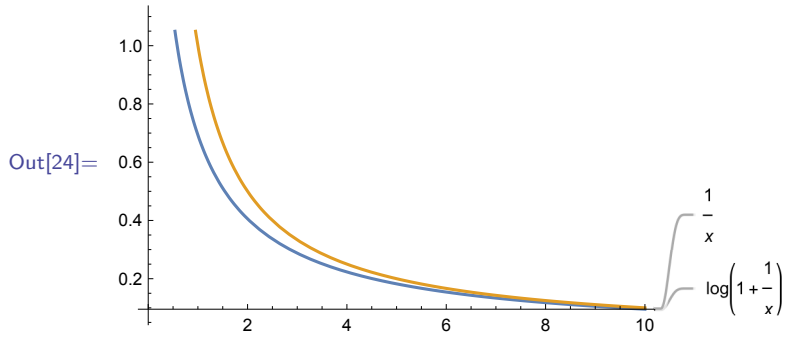
Luego $1 + \sqrt{5}a > \alpha^{n+2}$, es decir, $\log(1 + \sqrt{5}a) > (n+2)\log\alpha$. Usando que

$$\log(1+x) = \log x + \log(1 + 1/x)$$

y la desigualdad $\log(1+x) < x$ para todo $x > 0$, obtenemos que

$$(n+2)\log\alpha < \log(1+\sqrt{5}a) < \log(\sqrt{5}a) + \frac{1}{\sqrt{5}a}.$$

In[24]:=
```
Plot[{Log[1 + 1/x], 1/x}, {x, 0, 10}, PlotLabels -> "Expressions"]
```



Por lo tanto,

$$n \log \alpha < \log(\sqrt{5}a) - 2\log\alpha + \frac{1}{\sqrt{5}a}.$$

Despejando n se obtiene que

$$\begin{aligned} n &< \frac{\log(\sqrt{5}a)}{\log\alpha} - 2 + \frac{1}{\sqrt{5}a\log\alpha} \\ &= \frac{\log\sqrt{5}}{\log\alpha} + \frac{\log a}{\log\alpha} - 2 + \frac{1}{\sqrt{5}a\log\alpha} \\ &= c_2 + c_1 \log a - 2 + c_3 \frac{1}{a}. \end{aligned}$$

□

Corolario 2.1 El número de divisiones con residuo en el Algoritmo de Euclides es $O(\max\{\mathsf{Lon}(a), \mathsf{Lon}(b)\})$.

Teorema 2.11 La complejidad del algoritmo de Euclides para enteros a y b con $a \geq b$ es $O(\mathsf{Lon}(a)^3)$.

Demostración. Todas las operaciones en el Algoritmo de Euclides se hacen para enteros menores o iguales que a. Del Corolario 2.1 se tiene que el algoritmo de Euclides requiere $O(\mathsf{Lon}(a))$ divisiones con residuo. Cada una de estas divisiones tiene un costo del orden de $O(\mathsf{Lon}(a)\mathsf{Lon}(a))$. Luego el costo total es de $O(\mathsf{Lon}(a)^3)$ operaciones. □

Lo anterior se puede mejorar a $O(\mathsf{Lon}(a)\mathsf{Lon}(b))$ operaciones de bits, ya que a medida que se aplica el algoritmo de Euclides se van reduciendo el tamaño de las entradas. Para demostrar esto necesitamos el siguiente lema.

Lema 2.2 Sean $a, b \in \mathbb{Z}$ con $a \geq b > 0$, y sea q el cociente de dividir a entre b. Se tiene que $\mathsf{Lon}(a) - \mathsf{Lon}(b) - 1 \leq \mathsf{Lon}(q) \leq \mathsf{Lon}(a) - \mathsf{Lon}(b) + 1$.

Teorema 2.12 La complejidad del algoritmo de Euclides para enteros a y b es

$$O(\text{Lon}(a) \cdot \text{Lon}(b)).$$

Demostración. Supongamos que $a \geq b > 0$ y que $t = \text{Euc}(a, b)$. Por el Teorema 2.6 y el lema anterior se tiene que la complejidad del algoritmo de Euclides está dada por

$$\begin{aligned} \overbrace{\sum_{i=1}^{t} \text{Lon}(r_i)\text{Lon}(q_i)}^{\text{complejidad en cada división}} &\leq \text{Lon}(b) \sum_{i=1}^{t} \text{Lon}(q_i) \\ &\leq \text{Lon}(b) \sum_{i=1}^{t} (\text{Lon}(r_{i-1}) - \text{Lon}(r_i) + 1) \quad \text{(por el Lema 2.2)} \\ &= \text{Lon}(b)(\text{Lon}(a) - \text{Lon}(r_t) + t) \\ &\leq \text{Lon}(b)(\text{Lon}(a) + t). \end{aligned}$$

Como $t = O(\text{Lon}(a))$, entonces la complejidad es $O(\text{Lon}(a) \cdot \text{Lon}(b))$. □

Definición 2.5 Si a y b son enteros no ambos iguales a cero tales que m.c.d.$(a, b) = 1$, decimos que a y b son primos relativos.

Teniendo en cuenta que (Problema de Basilea)

```
In[25]:= Sum[1/i^2, {i, 1, Infinity}]
```

$$\text{Out[25]=} \quad \frac{\pi^2}{6}$$

es posible demostrar el siguiente resultado debido a Dirichlet del cual daremos un bosquejo de la demostración.

Teorema 2.13 (Dirichlet) Sean $a, b \in \mathbb{Z}$ elegidos aleatoriamente. La probabilidad de que a y b sean primos relativos es $6/\pi^2 \approx 0.60793\ldots$. Es decir, la probabilidad de que una pareja de enteros sean primos relativos es del 60 %.

Demostración. Se presenta un esquema de la prueba (hay detalles sobre la aleatoriedad a los que no entraremos). Sea $p = \text{Prob}(\text{m.c.d.}(a, b) = 1)$. Tenemos que $d = \text{m.c.d.}(a, b)$ si y sólo si

1. $d|a$.
2. $d|b$.
3. $(a/d, b/d) = 1$.

La probabilidad de que $d|a$ es $1/d$ (cada d-ésimo entero hay un divisor). Esta es la misma probabilidad para que $d|b$. La probabilidad de la tercer propiedad es p. Así que

$$\text{Prob}(\text{m.c.d.}(a,b)=d)=\frac{1}{d}\cdot\frac{1}{d}\cdot p=\frac{p}{d^2}.$$

Como

$$1=\sum_{d\geq 1}\text{Prob}(\text{m.c.d.}(a,b)=d)=p\left(1+\frac{1}{2^2}+\frac{1}{3^2}+\cdots\right),$$

entonces $p=6/\pi^2$. □

```
In[26]:= L = GCD[RandomInteger[10^10, 100], RandomInteger[10^10, 100]]

Out[26]= {1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 2, 1, 2, 9, 4, 1, 6, 1, 1,
          1, 3, 1, 1, 1, 4, 2, 1, 2, 1, 1, 2, 4, 2, 1, 3, 4, 4, 1,
          1, 7, 1, 3, 1, 1, 1, 1, 1, 2, 1, 1, 1, 1, 1, 2, 1, 1, 1,
          1, 2, 1, 1, 1, 1, 4, 2, 1, 1, 2, 1, 1, 1, 1, 2, 3, 1, 1,
          2, 1, 1, 1, 5, 1, 1, 1, 3, 1, 1, 2, 1, 1, 1, 4, 2, 1, 1,
          1, 1, 2, 1, 1}

In[27]:= N[Count[L, 1]/Length[L]]

Out[27]= 0.68
```

Sean $a_1,a_2,\ldots,a_k\in\mathbb{Z}$ (no todos nulos). El m.c.d.de $a_1,\ldots,a_k$ es el mayor entero d tal que $d|a_i$, para $i=1,\ldots,k$. Se puede demostrar que

$$d=\text{m.c.d.}(a_1,\ldots,a_k)=\text{m.c.d.}(\text{m.c.d.}(a_1,a_2),a_3,\ldots,a_k).$$

Como $\text{m.c.d.}(1,a_2,\ldots,a_k)=1$, entonces es *muy* probable que $k>1$ enteros escogidos aleatoriamente sean primos relativos.