Mini Apunte matemáticas discretas

Bernardo Subercaseaux

Mayo 2016

1 Introducción

El presente documento se propone presentar una ayuda a los alumnos y alumnas del curso de matemáticas discretas de la FCFM, a través de un esquema resumido de la materia, algunos ejercicios resueltos y otros tantos propuestos. Favor avisar cualquier error a bernardosubercaseaux@gmail.com

Es en la matemática donde el espíritu encuentra lo que más ansía : continuidad y perseverancia - Anatole France

2 Inducción y estructuras inductivas

2.1 Inducción débil

Sea P una propiedad cualquiera con dominio en los naturales, esto es, un predicado unario, que dado un número natural expresa algo, que puede ser cierto o falso. Por ejemplo:

$$P(n) \leftrightarrow n \text{ es primo} \leftrightarrow (\forall i \in \mathbb{N})\ i \mid n \rightarrow (i = 1 \vee i = n) \tag{1}$$

$$Q(n) \leftrightarrow 5 \mid n^5 - n \tag{2}$$

Mientras que la propiedad (1) es cierta para **algunos** naturales, la propiedad (2) es cierta para **todos**.

Indudablemente suena fuerte decir que algo es cierto para **todos** los números naturales, que es un conjunto infinito. Note que ningún programa es capaz de verificar esto. **Inducción** es una herramienta para este tipo de demostraciones, y como se verá, para mucho más.

El principio de inducción dice que si una propiedad es cierta para algún número natural n_o, y el que sea cierta en un número implica que lo es en el siguiente, se tiene entonces que será cierta en $n_0, n_0 + 1, (n_0 + 1) + 1, ...$ y así finalmente en todos los naturales mayores o iguales a n_0. Formalmente:

$$P(n_0) \wedge (P(n) \rightarrow P(n+1)\ \forall n \in \mathbb{N}) \leftrightarrow (\forall n \in \mathbb{N}, n \geq n_o)\ P(n) \tag{3}$$

Utilicemos como primer ejemplo la propiedad anterior. Que sea cierta para $n_0 = 0$ y que $Q(n) \rightarrow Q(n+1)$ implicaría, por principio de inducción, que la propiedad es cierta para todo natural. Por lo tanto, la labor se reduce a probar esas dos cosas.

$$Q(n_0 = 0) \leftrightarrow 5|0^5 - 0 \leftrightarrow 5|0 \leftrightarrow \textit{Verdadero} \tag{4}$$

A esto se le conoce como caso base. Por otra parte, está el caso inductivo, donde se utiliza $Q(n)$ como hipótesis para demostrar $Q(n+1)$.

En este caso, utilizaremos $5|n^5 - n$ para demostrar que $5|(n+1)^5 - (n+1)$. Pero desarrollando lo que queremos demostrar, se ve como $5|n^5 + 5n^4 + 10n^3 + 10n^2 + 5n - n$ pero sabemos que la suma de números divisibles por 5 es divisible por 5, entonces lo que queremos demostrar es $5|5(n^4 + 2n^3 + 2n^2 + 1) + (n^5 - n)$ donde el término de la izquierda es directamente divisible por 5, y el de la derecha es nuestra hipótesis inductiva, por lo tanto, acabamos de probar que asumiendo $5|n^5 - n$ se tiene $5|(n+1)^5 - (n+1)$, con eso, por principio de inducción, la propiedad es cierta para todos los naturales.

2.2 Inducción fuerte

Muy similar a lo visto anteriormente, consiste en probar un caso base *manualmente*, y luego probar un caso inductivo. En este caso, el caso inductivo requiere demostrar que la propiedad, digamos P, se cumple para $P(n+1)$ asumiendo, no solo que se cumple para $P(n)$ sino también para todos los naturales anteriores, hasta llegar al caso base.

Formalmente:

$$P(n_0) \wedge (\forall k, n_0 \leq k \leq n,\ P(k)) \ \rightarrow \ P(n+1) \leftrightarrow P(n)\ \forall n \geq n_0 \tag{5}$$

Como puede notarse, la única diferencia es que ahora la hipótesis es más $fuerte$ puesto que se asume que la propiedad es cierta para más valores, de ahí el nombre.

Típicamente se usa cuándo el caso inductivo depende de más de un valor anterior. Por ejemplo, los números de Fibonacci dependen de dos números anteriores, así que en general es conveniente usar sobre ellos inducción fuerte.

Ejemplo:

$$f_n \leq 2^n \quad \forall n \in \mathbb{N} \tag{6}$$

Dado que la propiedad se cumple para $n_0 = 0$, podemos tomar ese como el caso base. Supongamos ahora que la propiedad es cierta para todo k menor o igual que n, y demostremos ahora que se cumple para $n+1$.

Se tiene entonces por hipótesis:

$$f_{n-1} \leq 2^{n-1} \wedge f_n \leq 2^n \tag{7}$$

Pero la identidad fundamental de los números de fibonacci (típicamente se ocupa esta para cualquier demostración asociada a estos números).

Entonces $f_{n-1}+f_n = f_{n+1}$, y usando (7), $f_{n-1}+f_n \leq 2^{n-1}+2^n \leq 2^n+2^n \leq 2^{n+1}$ con lo que se tiene $f_{n+1} \leq 2^{n+1}$ y se concluye.

2.3 Principio del buen orden

Dice que cualquier subconjunto no vacío de $\mathbb{N}$ posee un menor elemento. Típicamente se usa de la siguiente manera:

1. Se desea probar una propiedad P para todo los naturales.
2. Se define el conjunto $S = \{n \in \mathbb{N} \mid n \textit{ no cumple } P\}$
3. Si $S = \phi$ estaríamos listos, entonces asumiremos S no vacío.
4. Como asumimos S es un subconjunto de los naturales no vacío, debe tener un menor elemento, digamos r.
5. Buscamos alguna forma de deducir que hay un elemento en S menor que r, lo que contradice que r sea menor. Así, la hipótesis de que S es no vacío, tiene que ser falsa, y por lo tanto se tiene que la propiedad es cierta.

Ejemplo:

Todo natural puede ser representado como suma de potencias de 2 distintas.

Demostración: Sea S el subconjunto de $\mathbb{N}$ de números que no pueden ser representados como suma de potencias de 2 distintas. Asumamos S no vacío, y sea entonces, en virtud del PBO, $r = min(S)$. Sea k la mayor potencia de 2 menor que r. (k siempre existe, piense un poco por qué). Ahora consideremos $p = r-k$. Claramente $p < r$. Supongamos p puede ser representado como suma de potencias de 2 distintas. Entonces k no puede estar en su representación, puesto que si lo estuviese, es decir, si $p = k + x$, dado que $r = p + k$, entonces $r = 2k+x$, pero $2k$ es una potencia de 2 más grande que k, lo que contradice que k sea la mayor potencia de 2 menor que r. Así, k no está en la representación de p, y por tanto $p + k = r$ es una suma de potencias de 2 distintas, lo que contradice que $r \in S$. Así la asumisión de que p puede ser representada como suma de potencias de 2 distintas tiene que ser falsa. Luego $p \in S$, pero como vimos, $p < r$, lo que contradice $r = min(S)$. Dado que exploramos todos los casos, y en todos ellos hay contradicción, la asumición inicial de que S es no vacío tiene que ser falsa, y por tanto todo natural puede ser representado como suma de potencias de 2 distintas.

Teorema: Los tres principios anteriores son equivalentes entre sí. Demostración propuesta como ejercicio, puede revisar la pauta del control del año 2013 para corroborar su solución.

2.4 Inducción Estructural

Una de las grandes gracias de la inducción es que sirve para mucho más que probar propiedades sobre los naturales, en realidad sirve para probar propiedades en conjuntos de *naturaleza inductiva*. Esto es, conjuntos que se definen a partir de un caso base, y donde los distintos elementos se componen de elementos *menores* relacionados. (Nótese que *menores* se refiere a un orden parcial arbitrario, propio del conjunto). Por ejemplo, el conjunto de los polígonos puede definirse inductivamente. El caso base son los triángulos, y luego puede notarse que cualquier polígono es la unión de un polígono y un triángulo. Esto se conoce como *triangulabilidad*. Y tiene el beneficio de que nos permite probar propiedades sobre todos los polígonos, como el Teorema de Pyck, demostrado en la auxiliar 5 o 6. De todos modos, un ejemplo que vale la pena mencionar es el de las palabras, que típicamente se definen inductivamente, y son de especial interés en computación. (En el curso de teoría de la computación se usará a diario esta definición estructural, por lo que vale la pena comprenderla bien).

Queremos pues, definir el conjunto de las palabras P, el caso base, la palabra más simple de todas, es la palabra vacía, que denotaremos ϵ. Además, las palabras están compuestas de caractéres, esto es, elementos en un alfabeto, digamos Σ. Con esto podemos definir:

1. $\epsilon \in P$
2. $w \in P \wedge \sigma \in \Sigma \rightarrow w\sigma \in P$

Así, por ejemplo, si $\Sigma = \{a, b, c\}$, entonces *abbcba* pertenece a P.

Definamos, para hacer un ejemplo, el lenguaje de las palabras de largo par. Para esto, haremos una definición estructuralmente inductiva, y luego probaremos estructuralmente que nuestra definición es correcta. ¿Cómo hacerlo? Bueno, partamos por el caso base, ¿Cuál es la palabra más simple? La palabra vacía. ¿Queremos que pertenezca a nuestro lenguaje? Sí, puesto que tiene largo 0 (par), por lo tanto, ya tenemos un caso base. ¿Que podemos hacer para asegurar que tendremos solamente palabras de largo par? Bueno, podemos hacer algo parecido a lo que hicimos antes, solo que en vez de agregar un solo caracter, agregamos dos cada vez, dando lugar a lo siguiente.

1. $\epsilon \in L$
2. $w \in L \wedge \sigma, \alpha \in \Sigma \rightarrow w\sigma\alpha \in L$

Ahora puede darse una definición inductiva del largo de una palabra, y demostrar así que toda palabra de L tiene largo par. Esto queda propuesto como ejercicio.

2.5 Algoritmos recursivos o basados en inducción

Uno de los usos más importantes y prácticos de las definiciones y argumentos inductivos es la posibilidad de construir algoritmos basados en ellos. En clases vieron el algoritmo de Kadane para la máxima suma de un subarreglo, por ejemplo. ¡Su correctitud la pueden demostrar inductivamente!

Demos otro ejemplo, supongamos que tenemos una palabra y queremos dividirla en la menor cantidad de trozos posibles, de modo que cada trozo sea un palíndromo. (Repase la materia dando una definición estructural de los palíndromos). Por ejemplo, la palabra *dificil* puede descomponerse en *d/ifi/c/i/l* dando un total de 5 trozos, cada uno palindrómico. Para esto definiremos $F(w, i, j)$ como la mínima cantidad de cortes necesarios para dividir la palabra w entre los indices i y j en palindromos. Entonces, si la palabra w entre los indices i y j es ya un palíndromo $F(W, i, j) = 0$ (Puesto que se necesitarían 0 cortes).. Así la respuesta final será $F(w, 0, |w|) + 1$ (puesto que si se requieren n cortes, resultan $n+1$ trozos). Plantearemos una solución en $O(n^3)$ que puede optimizarse fácilmente a $O(n^2)$.

$$F(w,i,j) = \begin{cases} 0 & \textit{si } w_i w_{i+1} ... w_j \textit{ es un palindromo} \\ \min_{k \leq j-1} \Big(1 + F(w,i,k) + F(w,k+1,j) \Big) & \textit{si no} \end{cases} \tag{8}$$

Recomiendo detenerse un poco a pensar en porque esta solución es correcta. A continuación lo demostraremos utilizando el PBO. Sea w una palabra arbitraria. S el conjunto de naturales n tales que existen dos indicies i,j tales que $j - i + 1 = n$ (así n representa el tamaño del intervalo entre i y j) y tales que $F(w, i, j)$ no es el mínimo número de cortes necesarios para dividir el la palabra en palíndromos en ese intervalo. Si S es vacío significa que F es correcto. Asumamos entonces que S es no vacío, por tanto tiene un menor elemento.Sea $m = min(S)$ y sean a, b los primeros indices tales que $b - a + 1 = m$ y la función falla en $F(w, a, b)$. Claramente $w_a w_{a+1} ... w_b$ no es un palíndromo, puesto que en ese caso la función retornaría 0, que efectivamente es el mínimo número de cortes necesarios. Dado que no es un palíndromo, requiere al menos 1 corte. Sea k la posición de alguno de los cortes necesarios (como se requiere al menos 1, esto está bien definido). $(i \leq k \leq j)$. Ahora bien, se debe asegurar que todas las partes sean palindrómicas, por lo tanto se debe dividir $w_a w_{a+1} ... wk$ y $w_{k+1} w_{k+2} ... w_b$ en partes palindrómicas. Ahora bien, F calcula bien el mínimo número de cortes en $F(w, a, k)$ puesto que $k - a + 1 < m$, y habíamos supuesto m menor elemento de S. así mismo, F calcula bien el mínimo número de cortes en $F(w, k + 1, b)$ puesto que el tamaño de ese intervalo es menor que el del mínimo en que falla, y por tanto si fallase en ese tendríamos una contradicción. Ahora bien, si se elije efectivamente cortar en ese k, entonces la respuesta sería correctamente 1 (el corte realizado en k) $+F(w, a, k) + F(w, k + 1, b)$. Ahora

bien, dado que existe algún k en que hay que realizar el corte, probando todos los valores de k del intervalo, y tomando la menor respuesta entre ellos, se tiene la respuesta que es efectivamente mínima. Note que se podría haber realizado el mismo argumento con inducción fuerte, e incluso débil, puesto que son equivalentes.

3 Conteo y probabilidades

3.1 Principios de conteo y probabilidades (suma y producto)

Si existen m formas de efectuar una acción, y n formas de efectuar otra, y son independientes, entonces hay $m \times n$ formas de efectuar la primera acción **y** la segunda. Así mismo, hay $m+n$ formas de efectuar la primera **o** la segunda. Por ejemplo, si hay tres películas posibles de ver (A,B y C), y 2 restaurantes a los que ir (P,Q) entonces hay 6 "panoramas" distintos llendo a una película y a un restaurante. En efecto {(A,P),(A,Q),(B,P),(B,Q),(C,P),(C,Q)}. Hay entonces 5 "panoramas" distintos llendo a una película o a un restaurante {A,B,C,P,Q}. Equivalentemente, en el contexto de probabilidades, dados dos fenómenos independientes, la probabilidad de que ocurran ambos, es decir, el primero **y** el segundo, es el producto de las probabilidades individuales. Análogamente, la probabilidad de que ocurra una **o** el otro es la suma de las probabilidades individuales.

En el caso en que los eventos son dependientes, es decir, tienen intersección, se debe aplicar **inclusión-exclusión** en el contexto de conteo y combinatoria y **probabilidad condicional/ Teorema de Bayes** en el contexto de probabilidades. Estos pueden ser consultados trivialmente en internet y no vale demasiado la pena detenerse al respecto aquí.

Un caso interesante que mezcla combinatoria y probabilidades es el **experimento de Bernoulli**,en distribuición binomial, (estos nombres dan lo mismo) consiste en realizar un experimento binario n veces, y analizar cuás es la probabilidad de que tenga éxito k veces. Suponiendo que la probabilidad de éxito individual es p. La fórmula que lo modela es:

$$\binom{n}{k} p^k (1-p)^{n-k} \tag{9}$$

Lo que podría llamar la atención es el combinatorio, ¿Por qué está ahí? La razón es sencilla, $p^k(1-p)^{n-k}$ es la probabilidad de obtener una secuencia de k éxitos consecutivos y luego $n-k$ fallos. Ahora bien, en general no nos interesa el orden en que se tienen los éxitos/fracasos. Por lo que nos interesan todas las secuencias que tienen k éxitos, independiente de que sean consecutivos o no. Por lo que tenemos que multiplicar por la cantidad de formas de ordenar esos éxitos. ¿Cuántas formas hay de ordenar los k éxitos? Bueno, podemos notar que

es equivalente a elegir un subconjunto de $[n]$ de tamaño k, y que los éxitos sean los experimentos de indices en ese subconjunto. Por ejemplo un subconjunto de [5] de tamaño 3 es {1,3,4} que equivale a la secuencia EFEEF, donde las E son éxitos y las F fracasos. Y sabemos que hay $\binom{n}{k}$ subconjuntos de [n] de tamaño k.

3.2 Esperanza y cálculo de la esperanza

Es la clásica herramienta que se utiliza para calcular el costo esperado de un algoritmo, es decir, lo que en promedio esperamos que tarde/utilice. Se describe, en un espacio discreto, con una variable aleatoria discreta X que toma valores $x_1, x_2, ..., x_n$ según :

$$E(x) = \sum_{i=1}^{n} x_i \times Pr(x_i) \tag{10}$$

Es una función claramente lineal, por herencia de la sumatoria/integral, que es lineal. Un buen ejemplo de uso está en el control 2 del semestre de otoño 2015.