Cambios de base

(muy muy en fase de realización)

Elvira Mayordomo
MOISES, Valladolid

septiembre 2004

Predicción de secuencias

- Un predictor de secuencias sobre {0,1,2}:

π: $\{0,1,2\}^* \times \{0,1,2\} \to [0,1]$

$\pi(w, a)$ es la probabilidad estimada de que w continúe con el símbolo a

$$\pi(w,0)+ \pi(w,1)+ \pi(w,2)=1$$

Predicción de secuencias

- El **error** del predictor π si el siguiente símbolo de w **es a** es $-\log_3(\pi(w,a))$

  Si $\pi(w,a)=1$ error 0

  Si $\pi(w,a)=1/3$ error 1

- Error acumulado a lo largo de w:

  $L_\pi(w)= \sum -\log_3(\pi(w[1..i-1],w[i]))$

Predicción de secuencias

- Error asintótico en una secuencia infinita x:

$$L_\pi(x) = \liminf_n L_\pi(x[1..n]) / n$$

Predicción de secuencias

- El predictor asocia una probabilidad a cada secuencia finita

$$p(w) = \prod \pi(w[1..i-1],w[i]))$$

$$\pi(w,a)= p(wa) / p(w)$$

- Error acumulado a lo largo de w:

$$L_\pi(w)= \sum -\log_3(\pi(w[1..i-1],w[i]))$$

$$= -\log_3(p(w))$$

Predicción de números

- ¿Qué pasa si las secuencias a predecir son representaciones de números?

  0.0211122.....

- ¿Son igual de predecibles en cualquier base?

Respuesta

- Depende de la pregunta
- Si se trata de buscar el mejor algoritmo que predice dentro de una cota de recursos (tiempo, espacio, ...)

    entonces a veces sí, a veces no

Recursos mínimos (memoria finita)

- Consideremos números en base 3 en los que sólo aparezcan los dígitos 0 y 1
- Un predictor que diga que nunca aparece el 2 tiene un error bastante bajo:

$$\Pi(w,0)=\Pi(w,1)=1/2 \qquad \Pi(w,2)=0$$

$$L_\pi(x[1..n])/n = \log_3 2$$

Recursos mínimos (memoria finita)

- Consideremos números en base 3 en los que sólo aparezcan los dígitos 0 y 1
- Si los traducimos a base 2 la mayoría son secuencias normales
  - Por tanto para cualquier predictor con memoria finita el error asintótico $L_{\pi}(x)$ es 1

Recursos mínimos (memoria finita)

- Consecuencia de lo anterior es que hay muchos números que en base 3 se pueden comprimir con autómatas finitos y en base 2 no
  - Porque compresión=predicción
- Como son muchos también se puede concluir lo mismo del algoritmo de Lempel Ziv
  - Porque el Lempel Ziv coincide con el mejor autómata para la mayoría de las secuencias

Sin cota de recursos

- Dado un predictor p para {0,1} quiero escribir uno para {0,1,2} que funcione tan bien como él al traducir cada secuencia con el cambio de base 2 a 3

- Dado n, hago uno que funcione para secuencias de hasta longitud n

Dibujito

Es decir

- Dado y en {0,1,2}* con |y|<= n
  cojo $m=(n+1)\log_2 3$

  $p_n'(y)= \sum p(w)$

¿Es eficiente?

- Si hay 3 intervalos seguidos dos tienen que ser w0, w1 y se suman

$$p_n(w0) + p_n(w1) = p_n(w)$$

Esto se repite y sólo pueden quedar $m=|w|=(n+1)\log_2 3$ intervalos

Eficiente

- A partir de tiempo polinómico, las predicciones son independientes de la base utilizada

Para dimensión

- Dimensión = predicción
- El conjunto de Cantor es en base 3

La alfombra de Sierpinsky

La alfombra de Sierpinsky

- En realidad las bases mezcladas no son equivalentes a las simples
- Pero la dimensión (=predicción) con bases mezcladas es una cota superior a la dimensión