

CADENAS DE MARKOV

CORTESÍA DE JIM SIMONS

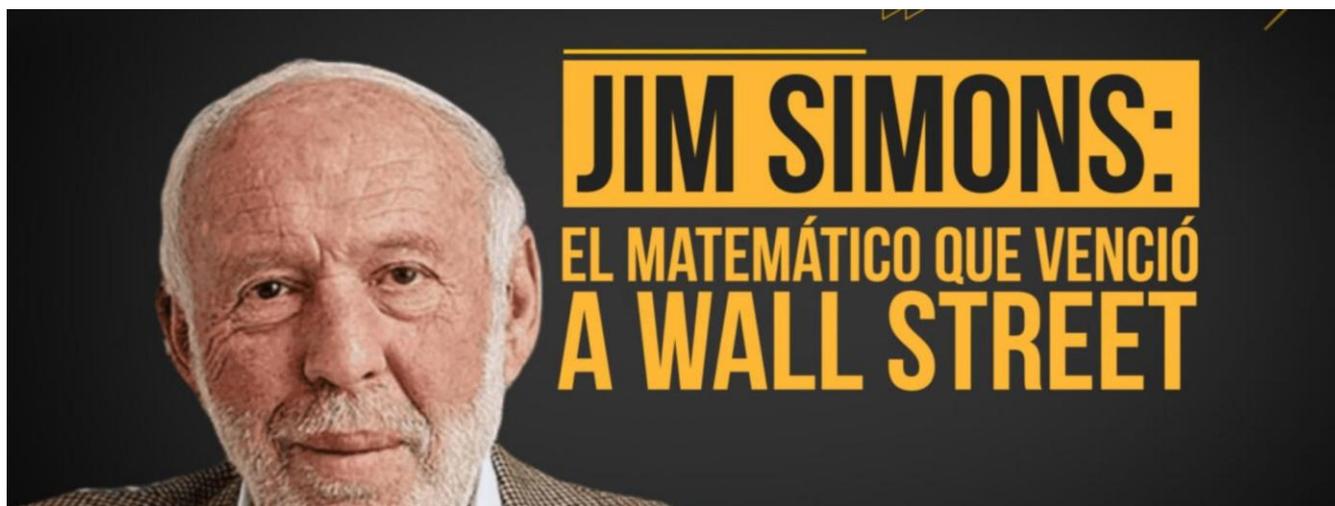
sábado, 27 de marzo de 2021

Horario para la semana Santa 2021: El sábado 3 de abril no habrá informe. En el resto no hay cambios.

Esta semana me entrevistaron en X-Trader, si quiere leerlo aquí está la entrevista: <https://www.x-trader.net/articulos/entrevistas-a-traders/entrevista-a-oscar-cagigas.html>

Es posible que haya oído hablar de Jim Simons (deba-jo), un matemático que trabajó descifrando códigos para la agencia nacional de seguridad americana durante la guerra de Vietnam aparte de contribuir a la teoría de cuerdas y a muchas otras áreas de la física.

Aunque Jim Simons es más conocido por su gestora Renaissance, creada en 1988 y que opera el Medallion Fund, un fondo que casi nunca pierde dinero. De hecho, el peor resultado en periodos de 5 años ha sido de un -0.5%. Invertir 1.000 dólares cuando comenzó el fondo ahora serían 13.8 millones de dólares. Impresionante...



Prohibida su distribución. La inversión en bolsa tiene riesgo. Utilice siempre Stop-Loss. Onda4 no se responsabiliza de las operaciones de sus seguidores. Onda4 puede utilizar este material en ofertas y/o promociones en su web.

210327

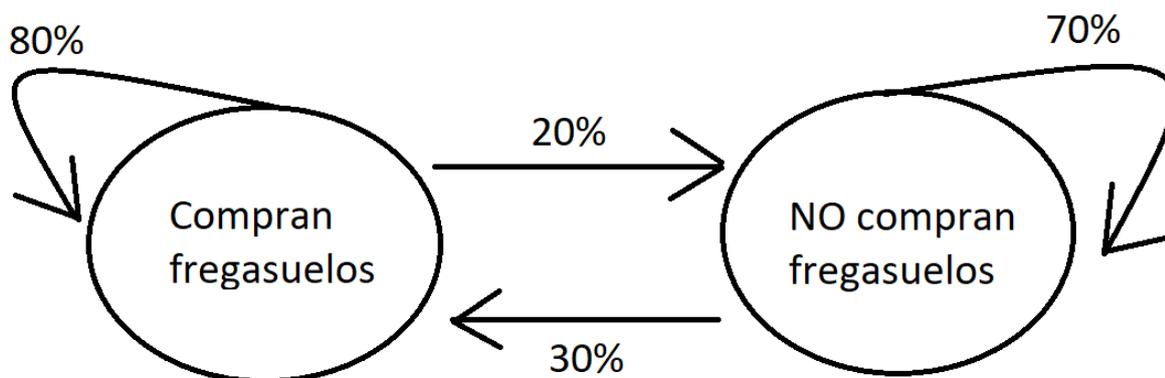
Pues gran parte de la operativa de Jim Simons está basada en Cadenas de Markov, aunque evidentemente a un nivel matemático muy avanzado. Hoy solamente rascaremos un poco en la superficie de este tema.

Una cadena de Markov es un proceso estadístico que establece cierta dependencia entre un estado y el estado anterior. Es decir, solo tiene memoria del estado precedente y lo demás no cuenta.

En finanzas esto es sencillo de entender. Supongamos que un estado significa retornos positivos. Y el otro estado, lógicamente, son retornos negativos. Así que una cadena de Markov podría decirnos cuál es la probabilidad de que hoy haya subidas si ayer hubo caídas. A continuación vamos a ver un ejemplo fuera de las finanzas que nos permitirá entenderlo:

El mercadona de VillaArriba sabe que el 20% de la gente que compra fregasuelos un mes no vuelve a comprarlo al mes siguiente. Y también saben que el 30% de los que no han comprado fregasuelos este mes sí que lo comprará el mes siguiente. A este Mercadona acuden 1000 personas este mes, y de esos 100 compraron fregasuelos. Cuántos lo comprarán el mes próximo? Y dentro de dos meses?

Aquí vemos los estados y sus probabilidades:



Dicho de otra forma: El 80% de los que compran fregasuelos este mes repiten su compra al mes siguiente, y el 70% de los que no compran fregasuelos este mes tampoco lo compran al mes siguiente. En realidad es un ejemplo bastante realista. En una pareja que viven juntos suele ser la misma persona la que compra este tipo de productos y la otra nunca o casi nunca se ocupa de ello.

Con la información aportada se puede construir la matriz de probabilidades

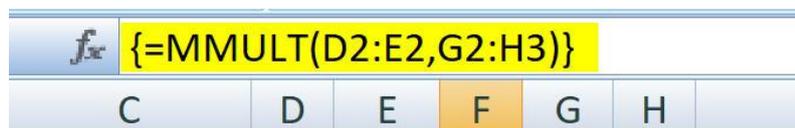
$$P^{(0)} = \begin{pmatrix} 0.8 & 0.2 \\ 0.3 & 0.7 \end{pmatrix}$$

Que si la multiplicamos por la matriz de personas que acuden al Mercadona pues nos tiene que dar el número de personas que el mes siguiente compra (C) o no compra (N). De la siguiente manera:

$$(C, N) = (100 \quad 900) \begin{pmatrix} 0.8 & 0.2 \\ 0.3 & 0.7 \end{pmatrix}$$

Para multiplicar estas matrices hay que recordarse del algebra, ya sabe, filas por columnas... Lo más sencillo es utilizar Excel.

En Excel usamos la función MMULT(), seleccionamos la fila y luego la matriz de probabilidades. El primer elemento sale 350. Si marcamos los dos elementos, pulsamos F2 y luego Ctrl+Shift+Enter pues nos sale el vector fila completo: (350, 650).



100	900	0.8	0.2
		0.3	0.7

Producto

350	650
-----	-----

La solución por tanto es que el primer mes comprarán fregasuelos 350 personas y 650 no lo comprarán.

Y el siguiente mes? Para ello hay que tener en cuenta que al igual que en probabilidades simples la probabilidad de una segunda ocurrencia de algo es la probabilidad inicial multiplicada por sí misma, así que para el segundo mes tenemos que calcular la matriz producto:

$$P^{(2)} = \begin{pmatrix} 0.8 & 0.2 \\ 0.3 & 0.7 \end{pmatrix} \begin{pmatrix} 0.8 & 0.2 \\ 0.3 & 0.7 \end{pmatrix} = \begin{pmatrix} 0.7 & 0.3 \\ 0.45 & 0.55 \end{pmatrix}$$

Para esto se puede hacer el mismo procedimiento en Excel pero ahora marcaríamos 4 casillas para tener el resultado anterior de 2x2.

De la misma manera calculamos el número de personas para el segundo mes:

$$(C, N) = (100 \quad 900) \begin{pmatrix} 0.7 & 0.3 \\ 0.45 & 0.55 \end{pmatrix} = (475, 525)$$

Es decir, el segundo mes 475 personas comprarán fre-gasuelos y 525 no lo comprarán.

Muy bien. Y cómo aplicamos todo esto a los mercados? Pues como decía en la página segunda podemos pensar en algo muy sencillo con dos estados. Nos cogemos los retornos del SP500 desde 2001; es decir, 20 años, y con la función JERARQUIA de Excel hacemos un ranking por retornos, de menor a mayor. Si dividimos el número de ranking entre el total de datos (5088) entonces tenemos el percentil que ocupa y así podemos calcular su estado de los dos posibles:

Date	S&P 500 retornos	Distribución	Estados
			2
03/01/2001	4.87%	99.37%	2
04/01/2001	-0.86%	17.06%	1
05/01/2001	-3.44%	1.55%	1
08/01/2001	0.41%	66.55%	2
09/01/2001	0.00%	44.60%	1
10/01/2001	1.42%	91.19%	2
11/01/2001	0.60%	73.98%	2
12/01/2001	-0.34%	29.34%	1
16/01/2001	0.42%	67.06%	2
17/01/2001	0.50%	70.32%	2
18/01/2001	1.13%	86.77%	2
19/01/2001	-0.59%	22.31%	1

Prohibida su distribución. La inversión en bolsa tiene riesgo. Utilice siempre Stop-Loss. Onda4 no se responsabiliza de las operaciones de sus seguidores. Onda4 puede utilizar este material en ofertas y/o promociones en su web.

En este punto es posible que vd piense que me estoy complicando demasiado, ya que solo serían retornos positivos o negativos, así que sobraría el ranking. Pero suceden dos cosas.

- No tiene por qué haber simetría (el cero no tiene que dividir los resultados en dos partes iguales)
- El método del ranking nos va a servir para aumentar el número de estados posteriormente

Ahora con una función CONTAR.SI ya podemos obtener la matriz de ocurrencias

Observac.	1	2
1	1255	1325
2	1325	1182

Si queremos las probabilidades solo hay que dividir entre el total de datos. Si el mercado fuera 100% aleatorio entonces esperaríamos $5088/4 = 1272$ ocurrencias de cada estado. Sin embargo vemos que es ligeramente más probable pasar del estado 1 al 2 (y viceversa) que de permanecer en el mismo estado.

En informes anteriores hemos visto que el mercado tiene una cierta tendencia a girarse al lado contrario del cierre anterior. Si se recuerda, el sistema 00:00 publicado en el informe del 9 de marzo estaba basado en esto.

En la tabla de la página anterior vemos como el primer día de trading de 2001 se sube un 4.87% (estado 2) y al día siguiente se tiene un retorno negativo perteneciente al estado 1. Este día muy alcista está prácticamente en lo más alto de los percentiles, con un 99.37%.

Y luego tenemos el 5 de enero que se cae un 3.44%, es el percentil 1.55% y por tanto el estado 1. Al día siguiente hay un cambio al estado 2.

Tal y como comentaba esto no está necesariamente centrado en el cero así que mirando entre los datos veo que el percentil del 50% es un retorno del 0.08%. Retornos mayores son el estado 2; y retornos menores, aunque sean positivos, son el estado 1.

Ahora que he explicado el procedimiento ya se puede extender a más estados, y no hace falta ni siquiera saber cuáles son los umbrales entre un estado y el siguiente ya que como acabo de explicar esto depende del total de los datos. El resultado para 4 estados es el siguiente:

Observ.	1	2	3	4
1	370	218	256	428
2	291	376	369	272
3	267	399	333	251
4	344	315	293	305

Que nos dice que es muy probable que después de un día muy bajista (estado 1) venga un día muy alcista (estado 4 con 428 ocurrencias). Y será raro que venga un día intermedio-bajista (estado 2 con solo 218 ocurrencias).

Así que para terminar me voy a construir un sistema de trading muy sencillo: Si hoy es estado 1 entonces mañana compro el SP500; es decir, del cierre de hoy al cierre de mañana. En la página siguiente vemos mi sistema de trading hecho en Excel. Si ayer hubo estado 1 entonces hoy se anota la ganancia del día pues se asume que es lo que se gana al hacer la compra (no se incluyen comisiones y se asume que operamos justo por el nominal). Por ejemplo, el 4 de enero hay un estado 1 y al día siguiente se anota una pérdida del 3.44% en el día. Ese mismo día también es estado 1 así que se repite la compra al día siguiente, esta vez con retorno del 0.41%. Aquí cada operación dura un día solo.

G5		=SI(F4=1,B5,0)					
	A	B	C	D	E	F	G
1	Date	S&P 500 returns	Distribución	Estados			Strategy return
2				2	3	4	
3	03/01/2001	4.87%	99.37%	2	3	4	
4	04/01/2001	-0.86%	17.06%	1	1	1	0.00%
5	05/01/2001	-3.44%	1.55%	1	1	1	-3.44%
6	08/01/2001	0.41%	66.55%	2	2	3	0.41%
7	09/01/2001	0.00%	44.60%	1	2	2	0.00%
8	10/01/2001	1.42%	91.19%	2	3	4	0.00%
9	11/01/2001	0.60%	73.98%	2	3	3	0.00%
10	12/01/2001	-0.34%	29.34%	1	1	2	0.00%
11	16/01/2001	0.42%	67.06%	2	3	3	0.00%
12	17/01/2001	0.50%	70.32%	2	3	3	0.00%
13	18/01/2001	1.13%	86.77%	2	3	4	0.00%

Y ahora ya puedo acumular los resultados de la columna que tiene la estrategia. Comienzo con 100.000 dólares y resulta la siguiente curva de capital:



Prohibida su distribución. La inversión en bolsa tiene riesgo. Utilice siempre Stop-Loss. Onda4 no se responsabiliza de las operaciones de sus seguidores. Onda4 puede utilizar este material en ofertas y/o promociones en su web.