



## LA IMPORTANCIA DE LAS BASES DE DATOS EN UN BACKTEST

# Los datos son nuestra materia prima

Los gestores que utilizamos el trading algorítmico como principal herramienta de gestión, tratamos de confeccionar carteras de sistemas que contengan estrategias con Esperanza Matemática Positiva y baja correlación entre ellas. Esto, junto con una buena gestión del capital es lo único necesario para obtener retornos positivos consistentemente. ¿Sencillo? Obviamente, no todo el monte es orégano, pero es indudable que un BackTest bien hecho es condición necesaria, aunque no suficiente para saber si estamos ante un sistema aprovechable para nuestra cartera. Sabemos que rentabilidades pasadas no garantizan futuras, pero un BackTest bien hecho es necesario para evaluar correctamente a nuestra estrategia y saber qué podemos esperar y qué no de una estrategia.

### ¿Qué es un Backtest bien hecho?

Un BackTest es importante, de acuerdo, pero ¿cómo hacemos un BackTest bien hecho? No es objeto de este artículo detallar el procedimiento para hacer



#### Sergi Sánchez Alvira

Sergi es Gestor de Esfera/Sersan Algorithmic, vehículo de inversión que opera con sistemas cuantitativos y automáticos de trading. ESFERA/SERSAN ALGORITHMIC es un compartimento de ESFERA, FI. Fundador de Sersan Sistemas, también experto en Trading algorítmico con años de experiencia en el desarrollo, testeo, evaluación y sobre todo, en la Gestión cuantitativa con sistemas automáticos de trading.

✉ [sergisanchez@esferacapital.es](mailto:sergisanchez@esferacapital.es)

un BackTest si no tratar sobre la materia prima de un BackTest, la base de datos. Además, existen muchos caminos o tipos de BackTest diferentes, dependiendo del objetivo perseguido por el mismo o incluso las preferencias personales. Lo que sí tienen en común todos ellos es la necesidad de usar una base de datos adecuada. Esto es imprescindible y por experiencia, es un aspecto al que muchas veces no se le presta demasiada atención. Simplemente se usan aquellos que tenemos en la plataforma de turno sin cuestionarnos ni verificar si los datos son correctos o cumplen los criterios necesarios para obtener resultados extrapolables. Si pretendemos obtener conclusiones realistas es imprescindible trabajar con datos de calidad y eso pasa por evaluar los datos antes de empezar un BackTest.

En términos generales podemos definir a un BackTest bien hecho como una simulación realizada sobre datos históricos de calidad, siendo la muestra estadísticamente significativa (que tenga suficientes trades) y

representativa del universo objeto de estudio (que cubra el máximo tipo de mercados posible) Debemos tratar de simular las condiciones que nos encontraremos operando en live, como pueden ser las comisiones o las desviaciones sobre los precios teóricos y los ejecutados (slippage) o que los datos tengan propiedades homogéneas, por ejemplo, que los cierres tengan lugar a la misma hora. Como no, deberemos realizar diversas pruebas de stress, etc. pero esto no es objeto de este artículo. Centrémonos en los datos, nuestra materia prima.

### Datos de calidad

Como en todo, hay feeds de datos en tiempo real mejores y peores y lo mismo ocurre con los datos históricos e igualmente con el software que tiene que procesar esa información, datos y reglas del sistema.

Estas tres partes son importantes para comprender lo que es una base de datos de calidad. Los datos históricos tienen que ser homogéneos en comportamiento y construcción a los datos live que recibimos en tiempo real. Y el procesamiento de los datos que hagamos en live debe ser el mismo que hacemos en el BackTest. Dicho de otra manera, el BackTest tiene que seguir las mismas reglas de evaluación que seguirá en live, algo que no todas las plataformas hacen, con lo que los resultados obtenidos en BackTest no serán reproducibles en live. Esto último es también muy descuidado o mejor dicho ignorado, simplemente ni nos cuestionamos si la forma de evaluar las reglas de nuestro sistema es igual en live que en BackTest y esto es sencillamente vital para poder valorar correctamente al sistema. Pues no lo es en muchas ocasiones, pero volveremos a ello más adelante. Sigamos ahora con los datos.

En efecto, la muestra debe ser representativa del universo objeto de estudio y también debe ser representativa del comportamiento actual del mercado. Es muy frecuente ver que en la mayoría de las plataformas o "feed" de datos, que los datos históricos recientes tienen buena calidad, sin demasiados gaps erróneos o ticks perdidos. En cambio, a media que vamos retrocediendo hacia atrás, la base de datos va perdiendo calidad y empiezan a aparecer errores visibles incluso a simple vista. Si no nos tomamos la molestia de revisar el histórico esto nos pasará por alto. El problema es que, estos errores no siempre son visibles a simple vista, por lo que necesitamos algún método para comprobar si los datos históricos son correctos o no. El mejor camino es contratar los datos a un proveedor externo que se encargue de realizar esta verificación diariamente. Así,

podremos comparar los datos de nuestra plataforma con los que hemos comprado e incluso usar estos datos para nuestros BackTest si la plataforma nos permite importarlos. La comprobación visual puede ser una tarea ardua pero necesaria; no obstante, puede hacerse también mediante un código que compare ambas bases de datos y pinte en un indicador la diferencia entre ellos o incluso hay algunas plataformas que incluyen herramientas para comparar bases de datos y nos generan un informe al respecto.

Y lo mismo aplica a los datos en tiempo real. Aunque aquí debería de haber poca diferencia entre proveedores de datos, la realidad es que en ocasiones la hay. Dependiendo del tipo de barras que usemos esto es más crítico o menos. El gráfico de ticks suele ser muy conflictivo ya que al construir las barras usando n número de ticks, si el feed de datos pierde algunos ticks todas las barras que se construyan durante la jornada con posterioridad a esa pérdida quedarán comprometidas. En cambio, en las barras construidas por tiempo, una pérdida de algunos ticks afectará a la barra en cuestión, pero no afectará a las posteriores. Por tanto, si queremos trabajar con ticks hay que prestar mucha atención en evaluar al proveedor de datos en tiempo real. También debemos estar atentos si usamos barras por tiempo. Una de las principales fuentes de errores es el estampado de la hora de las velas, lo que puede afectar de manera crítica a los cierres del gráfico.

Los datos en tiempo real son algo que también tenemos que verificar y, en todo caso, recordar que los datos de BackTest deben tener las mismas propiedades y comportamiento que los de live. Si los datos en live vemos que no son correctos comparados con una base de datos verificada como fiable, tenemos un problema. En ocasión un colega me comentó que la solución a este problema es optimizar en los datos de este mismo proveedor que estamos viendo que no difunde correctamente los datos. Él defendía que, aunque los datos estén mal, si optimizamos con ellos ya estamos teniendo en cuenta el error en la optimización y, por lo tanto, el sistema se ajustará al error. En mi opinión, esto es totalmente absurdo. Los datos reales, veraces, son solo unos. El mercado ayer cerró a un precio, hizo un máximo y un mínimo. Cualquier otro dato que no sea el real es fruto de una casualidad o mala praxis en el etiquetado de las barras, por ejemplo. Por tanto, cualquier regla extraída de esos datos es casual porque un error de difusión de datos no es estable. Un error es un error. Recordemos que un sistema explota una ventaja o ineficiencia, trata de capturar la señal del mercado que explota tratando de

evitar o filtrar el ruido. ¿Cómo podemos conseguir que el sistema detecte o aprenda de esa señal si lo estamos alimentando con información no veraz y por tanto no reproducible?

**Construcción de nuestra base de datos**

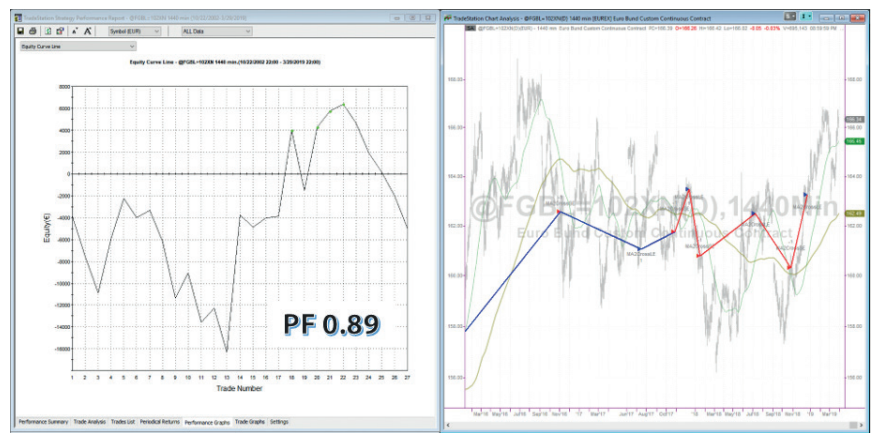
Ya sabemos que las bases de datos que queremos usar provienen de fuentes confiables y son veraces. Si trabajamos con sistemas que operen en barras diarias, probablemente nos interesará usar la barra intradiaria para construir la diaria. Esto es porque algunos proveedores de datos modifican el valor del cierre del futuro con posterioridad al cierre del mercado, para reflejar el precio de liquidación diaria oficial del futuro. En efecto, todos los futuros cierran oficialmente a un precio, que es el que se usa en la industria para liquidar las posiciones a Mark to Market. Aunque operemos en vela diaria, a nosotros este precio no nos aporta nada ya que no es fruto de la acción de la oferta y la demanda sobre el precio. De hecho, suele aparecer horas después de que el mercado haya cesado su actividad. Los proveedores que reflejan este precio en los gráficos suelen hacerlo en la base de datos diaria, pero no en la intradiaria, que solo refleja los precios intercambiados realmente en el mercado. Por tanto, si construimos la vela diaria con la vela intradiaria, evitaremos que el precio se modifique con posterioridad. Un día tiene 1440 minutos o si operamos los futuros sobre índices norteamericanos en horario regular, la sesión de 15:30 a 22:15 (hora española) dura 500 minutos. Por tanto, una vela de 500 minutos equivale a la vela diaria construida con la base intradiaria. Es algo que tenemos que verificar si operamos sistemas en velas o barras diarias. Ahora nos toca verificar si los datos están ajustados o no y decidir como los queremos configurar. Dependiendo si vamos a trabajar con futuros o con acciones tendremos problemáticas distintas, aunque

análogas. Podemos decir que las bases de datos sufren gaps súbitos que no son frutos de la acción del precio si por eventos externos a ésta.

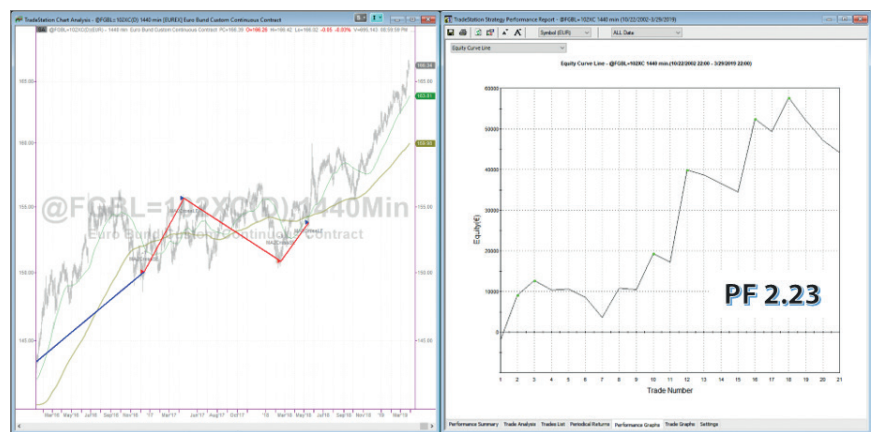
Las acciones suelen requerir ajustes por el pago de dividendos u otros ajustes sobre el capital como splits o contra splits.

Los futuros solo requieren ajuste si utilizamos el contrato continuo que se construye por la unión de diferentes vencimientos. Frecuentemente, en el momento de la unión existe una diferencia de precios entre los dos vencimientos provocados por diferentes factores como el tiempo hasta el vencimiento, el posible cobro de dividendos del subyacente entre otros. La mayoría de los proveedores de datos suelen incluir el símbolo continuo

**G1 Sistema Golden Cross aplicado al futuro del BUND sin ajustar en barras diarias desde 2002**



**G2 Sistema Golden Cross aplicado al futuro del BUND ajustado en barras diarias desde 2002**



En el Gráfico 1 y en el Gráfico 2, podemos ver el resultado de una sencilla estrategia llamada Golden Cross. En el Gráfico 1 la hemos aplicado al Gráfico del BUND sin ajustar y en el Gráfico 2 al Gráfico del BUND ajustado, con la misma configuración de ajuste que el Gráfico 3. En ambos casos en barras diarias desde finales de 2002. La estrategia abre largos a mercado en la apertura de la siguiente vela cuando la media de 50 cierra por encima de la de 200 y justo lo contrario para el lado corto. Es un sistema stop and reverse que siempre está en el mercado. Los resultados hablan por sí solos.

Fuente: TradeStation

## Los futuros solo requieren ajuste si utilizamos el contrato continuo que se construye por la unión de diferentes vencimientos.

para poder trabajar con una mayor cantidad de histórico ya que cada vencimiento tiene un periodo limitado de histórico, pero no todos los proveedores ajustan esta base de datos.

La necesidad de ajuste, tanto en un caso como en otro, proviene de los gaps que se generan cuando se paga un dividiendo, en el caso de las acciones, o cuando se unen vencimientos, en el caso de los futuros. Es cierto que hay posturas enfrentadas sobre si hay que ajustar o no las bases de datos. En mi opinión, si el objetivo que perseguimos es BackTestear un sistema o varios en esa base de datos, el ajuste es necesario. El motivo es que esos gaps no son fruto de la oferta y la demanda del precio, son exógenos. A nosotros no nos interesa si es relevante económicamente o no o si con eso alteramos los precios históricos. Lo que nosotros necesitamos es una base de datos homogénea que refleje la acción del precio fielmente, sin gaps artificiales que estropeen la relación de los precios. No nos importa tanto los valores absolutos de los precios si no su relación que, recordemos, debe ser representativa de lo que ocurrirá en live por lo que cualquier ajuste debe ser también reproducible en live.

Expliquemos un poco más en detalle la casuística de los futuros. Los futuros tienen vencimientos, trimestrales o mensuales, lo que hace que usualmente utilizemos símbolos continuos para hacer el Back-Test. Los símbolos continuos unen los distintos vencimientos uno detrás de otro, cambiando de uno a otro cada vez que un contrato expira o pocos días antes. El problema surge en el momento de enlazar estos vencimientos. Usualmente el vencimiento antiguo y el nuevo tienen una diferencia entre sus precios, lo que provocaría que el día en que se enlacen en el símbolo continuo, se produjera un gap o diferencia artificial que no ha ocurrido realmente en el mercado.

Tres variables influyen en el ajuste a realizar:

1. ¿Qué vencimientos usamos? Usualmente se usan todos los disponibles para construir el símbolo continuo, esto es, 4 si el vencimiento es trimestral, 12 si es mensual. No obstante, hay futuros que tienen particularidades que hacen que no se usen

todos. Por ejemplo, sobre el futuro del oro cotizan los vencimientos F, G, J, M, Q, V, X y Z<sup>(1)</sup>, pero para construir el continuo solo usamos los vencimientos G, J, M, Q y Z<sup>(1)</sup>, es decir, excluimos F, V y X<sup>(1)</sup>. Los motivos suelen tener su origen en algún fenómeno estacional del subyacente, por ejemplo, el periodo de recolección o la meteorología, que acaban siendo uso y costumbre para el mercado. Si no conocemos este dato podemos verificar cual es el que el mercado considera como el Front comparando los vencimientos individualmente, que será el que esté negociando más volumen y/o tenga mayor open interest.

2. ¿Qué método de roló utilizamos? El método más utilizado para cambiar de vencimiento es el volumen y/o el open interest. La regla general es que el gráfico continuo siempre contenga aquel vencimiento que está negociando más volumen o tiene mayor open interest, esto es, el que el mercado considera como Front o principal. Hay softwares específicos de construcción de bases de datos que ya incluyen un método automático que justamente hace eso, mantener en el símbolo continuo aquel que negocia más volumen. En la práctica, los participantes del mercado rolan sistemáticamente el mismo día por uso y costumbre, por lo que tras verificar en qué día un vencimiento pasa a hacer más volumen que el otro con relación a la expiración, podemos sistematizar el cambio de contrato usando como referencia esos n días antes del vencimiento o expiración. Por ejemplo, los futuros sobre índices norteamericanos rolan 6 días antes de la expiración, el del DAX, CAC o IBEX lo hace el día antes de la expiración y el del oro lo hace 2 días antes.
3. ¿Qué tipo de ajuste usaremos? Para ajustar la diferencia de precios que existe entre los dos vencimientos a enlazar, podemos hacerlo de dos maneras.
  - 3.1 La más común es el ajuste Backward. En este caso restamos a todos los precios pasados la diferencia entre los dos contratos en el momento del enlace. También puede hacerse dividiéndolo

(1): Las letras de los vencimientos de futuros corresponden a los siguientes meses: F: Enero - G: Febrero - H: Marzo - J: Abril - K: Mayo - M: Junio - N: Julio - Q: Agosto - U: Septiembre - V: Octubre - X: Noviembre - Z: Diciembre



### G3 Comparación entre el futuro del BUND ajustado y sin ajustar en velas semanales desde 2013



En el Gráfico 3 podemos ver el gráfico continuo sobre el futuro del BUND alemán en velas semanales. En la parte de arriba tenemos el símbolo sin ajustar y en la parte de abajo el ajustado usando todos los vencimientos, rolado dos días antes del vencimiento y mediante el método Backward restando la diferencia en valor absoluto. El BUND es uno de los futuros más afectados por la gran diferencia que suele haber entre los vencimientos. Por este motivo vemos que el de la parte de arriba parece que esté en lateral hace muchos años mientras que el ajustado no deja de hacer máximos históricos. El que refleja mejor la realidad económica, como se puede ver en el gráfico de los tipos alemanes a 10 años que está en valores negativos (es el inverso del BUND), es el ajustado.

Fuente: TradeStation

por una ratio generada con la diferencia entre los dos contratos en porcentaje. Este último método plantea el problema de que los precios ajustados pueden pasar a tener decimales que no tienen la base sin ajuste. Hay que pensar o bien en redondear la base de datos o en redondear los precios en el código del sistema. En todo caso mantendremos la relación que existía entre los precios antes del enlace, pero perdemos la referencia histórica. Es decir, puede ocurrir que el máximo del histórico antes del ajuste fuera 2000 puntos y tras el ajuste sea 1997. Si ajustamos para BackTest probablemente esto no nos importe si a cambio mantenemos la relación de los precios y evitamos los gaps falsos. Un posible problema que plantea este método es que en determinados futuros donde el ajuste a hacer es de un tamaño importante, puede ocurrir que los precios pasados acaben en números negativos lo que tendremos que gestionar también mediante código.

3.2 La otra forma de ajustar los datos es Forward. En este caso añadimos la diferencia entre los vencimientos a los precios futuros desde el instante del enlace. Así, mantendremos las referencias históricas pero los últimos precios que veamos en el gráfico, los más recientes, no serán los de mercado ya que tendrán añadido un plus para el ajuste, que vencimiento tras vencimiento puede llegar a ser importante. Es el usado en los índices Total Return. El método anterior, en cambio, siempre tendrá en la parte derecha del Gráfico, la más reciente, los precios reales de mercado, ya que los ajustes se han hecho a los precios pasados.

El caso de las acciones y/o ETFs es más sencillo al no tener distintos vencimientos por lo que solo aplica el último punto, el método de rolo, Backward o Forward.

Cuando se produce un pago de dividendos, ese importe es descontado del precio de la acción inmediatamente. De hecho, si tenemos las acciones en nuestra posesión,

nuestro estado contable no habrá variado ni 1 céntimo en términos brutos tras el cobro del dividendo; en términos netos sí, porque hacienda se quedará con su parte del dividendo cobrado. Si la acción cotiza a 20€ y paga 1€ de dividendo la acción empezará a cotizar la siguiente sesión a 19€ y nosotros tendremos en nuestra cuenta 1€ menos la retención de hacienda. Esta diferencia de 1€ en el precio origina un gap que no ha sido fruto de la oferta y de la demanda y es aquí donde nace la necesidad del ajuste. En este caso, ni tan solo tiene sentido económico ya que el inversor ha ingresado ese euro por lo que el gráfico que refleja mejor su situación económica es el ajustado, sin ese gap.

De hecho, ya hace muchos años que los principales mercados publican sus índices en versión Total Return (TR), esto es, el índice con la reinversión de los dividendos de los componentes del índice. Estos índices usan la metodología de ajuste Forward, por eso el precio de los índices TR es muy superior al índice. Estos índices son los Benchmarks adecuados para comparar una

inversión indexada como la mayoría de los fondos de inversión pasivos, algo que la mayoría de la industria no hace, algo que, en mi opinión, es una trampa al inversor. Dicho de otra forma, cualquier fondo que no bata a su índice de referencia, por ejemplo, el IBEX-35, por lo menos por los dividendos que paga ese índice, está haciéndolo peor que comprar directamente el índice. Compararse con los índices Total Return, que ya recogen este efecto, es lo más correcto económicamente y lo más ético moralmente ya que reflejan la realidad de un inversor que hubiera comprado las acciones que componen el IBEX-35 de manera indexada o que hubiera comprado un ETF sobre el IBEX-35, ya que si hubiera hecho esto también ingresaría los dividendos que pagan las acciones, como hace el índice Total Return o TR.

Se puede pensar que si trabajamos con gráficos intradiarios no hay necesidad de complicarse la vida con ajustes. Los días posteriores al vencimiento de un futuro y de su cambio de contrato en el símbolo continuo o del cobro de un dividendo en el caso de las acciones, cualquier indicador o cálculo realizado sobre datos históricos se verá afectado por ese "falso" gap del Gráfico continuo. Por tanto, si trabajamos con bases de datos no ajustadas, porque nuestra plataforma no lo permite, debemos contemplar este potencial problema en los datos y en los resultados obtenidos.

### La optimización

Hay que tener en cuenta que si optimizamos es especialmente importante el mantener una base de datos lo más veraz posible, es decir, lo más representativa de lo que ocurrió realmente en el mercado y de lo que probablemente ocurrirá en el futuro. Si no excluimos estos falsos gaps, condicionaremos al optimizador, especialmente si es genético, de manera artificial. Dificultaremos encontrar zonas robustas o peor aún, podemos incurrir en conclusiones erróneas pensando que una zona es robusta cuando no lo es.

El optimizador genético ayuda a encontrar zonas robustas por su lógica de funcionamiento. Este va

### G4 Comparación entre IBEX-35 e IBEX-35 TR en barras semanales desde el año 2000



En el Gráfico 4 tenemos un ejemplo sobre el IBEX-35, uno de los índices que más dividendos paga por lo que se ve muy afectado por este efecto en plazos largos. En el momento de la captura de la imagen el IBEX-35 está a 9240 puntos y el IBEX-35 TR está a 25959 puntos. Se aprecia claramente como este último ha superado los máximos que marcó en diciembre de 2007 mientras que el IBEX-35 está muy lejos de ese máximo que fue en los entornos de los 16000 puntos. Es decir, una inversión indexada al índice será muy parecida a IBEX-35 TR, excluyendo el pago de impuestos. De hecho, algunos mercados también difunden un índice que se llama Net Total Return (NTR) que recoge también el pago de impuestos.

Fuente: Visual Chart

buscando los sets que mejor valor obtiene de una determinada función objetivo y por mutación va mejorando la selección de los sets. Por definición, es más sencillo que encuentre zonas donde hay muchos valores buenos que donde hay un o unos valores buenos aislados. Las reglas evaluadas explotan una determinada ineficiencia en los precios y si estos son erróneos, como ya he explicado anteriormente, estamos tirando piedras sobre nuestro tejado, especialmente grandes si optimizamos. Es básico optimizar en bases de datos limpias, representativas y bien ajustadas. Los datos lo son todo, son nuestra materia prima. Obviamente, disponer de herramientas que nos permitan ajustar o no los datos o incluso configurar cómo y cuándo hacer el ajuste, siempre será la mejor opción.

A partir de aquí, deberíamos elegir muestra, designar los segmentos In Sample y Out Sample si queremos, configurar la optimización correctamente para que sea homogénea en evaluación al comportamiento live... pero todo esto sería objeto de otro artículo. Hoy quería centrarme en las bases de datos ya que por mi experiencia es un aspecto descuidado incluso por algunos profesionales del trading algorítmico.