



OPTIMIZACIÓN WALK-FORWARD V OPTIMIZACIÓN BRAC

Riesgos y evaluación de robustez

La optimización es un proceso controvertido dentro del trading algorítmico, tiene detractores y defensores por igual. La optimización de una estrategia es el proceso de probar una combinación de valores o inputs de la estrategia para encontrar la combinación que da los mejores resultados sobre unos datos históricos, utilizando una función objetivo especificada. Es un proceso intensivo de prueba error, que se aprovecha de las evoluciones tecnológicas que han mejorado mucho la velocidad de procesamiento, que va comprobando qué resultados hubieran obtenido en los datos históricos una determinada combinación de parámetros y nos los recopila todos. Hay varios tipos de optimización y diversos procedimientos para optimizar, pero no son el objetivo de este artículo.

Optimización: tan potente como peligrosa

Decía que tiene sus detractores y defensores. Los primeros alegan que es muy fácil caer en lo que llamamos sobre-optimización por lo que prefieren no usarla. Esto es, que mediante la optimización obtengamos unos parámetros que satisfagan los targets que hemos prefijado, pero que estos no sean fruto de una pauta estadística sólida o de una ventaja, sino de una casualidad estadística. Así, cuando nos pongamos a operar con ellos, su resultado será mucho peor del que cabría esperar. Tengamos en cuenta que nosotros usamos lo que se conoce como estadística inferencial o inductiva, es decir, analizamos una parte de una muestra de datos y deducimos unas conclusiones que extrapolamos a la totalidad de la muestra. Mediante el proceso de optimización elegimos unos parámetros que, por ejemplo, maximizan el ratio de Sharpe en el histórico, es decir, que han funcionado muy bien en esos datos históricos. El problema es que lo que nosotros queremos saber realmente es si esos parámetros

y nuestro sistema ganará dinero en el futuro. Que en el pasado haya funcionado muy bien es condición necesaria, pero no suficiente. Aquí es donde los detractores basan la mayoría de sus argumentos. Si realmente el resultado de la optimización no solo no garantiza ganar dinero en el futuro, sino que puede confundirme o engañarme, al indicarme que una combinación de parámetros es muy buena, pero luego en la operativa real no acaba siéndolo ¿por qué usarla? La sobre-optimización o curve fitting es un riesgo real, es fácil ajustarse a la curva y debemos trabajar para minimizar su riesgo lo máximo posible.

Los defensores de la optimización afirman que existen diversas herramientas para evaluar la robustez, así como pautas de diseño que pueden fomentarla durante el desarrollo. También afirman que usar una determinada combinación de parámetros por que sí, sin adaptarlo a cada activo o timeframe, es perder una ventaja muy grande. Es cierto que es complicado medir la robustez. Hay herramientas como el Cluster Walk-Forward Analysis que veremos hoy, pero a mí me gusta definir a la robustez como la capacidad para mantener en el futuro los datos de performance obtenidos con datos históricos optimizados o backtesteados. Es decir, la certeza de robustez solo la tendremos cuando estemos ya operando, tarde, por tanto. Las herramientas que podemos usar para medir la robustez, si las usamos bien, aumentan mucho la probabilidad de que funcionen en el futuro, pero la certeza no existe nunca a priori. En realidad, tanto detractores



Sergi Sanchez

Gestor de Esfera Sersan Algorithmic F.I. Director Ejecutivo de Sersan Sistemas. Experto en Trading algorítmico con años de experiencia en el desarrollo, testeo, evaluación y, sobre todo, en la Gestión cuantitativa con sistemas.

✉ sersan@sersansistemas.com

como defensores tienen su razón y por eso es aconsejable conocer bien los riesgos que tiene la optimización, las herramientas que existen para hacerla y para evaluar la robustez y a partir de ahí usarla o no dependiendo del tipo de sistema, del tipo de datos, etc.

Ciencia aplicada al trading

Dentro de la producción o creación de sistemas, la optimización la podemos ubicar en la parte de evaluación de estrategias que es aquella parte en la que tratamos de usar la ciencia para demostrar que la idea de trading,

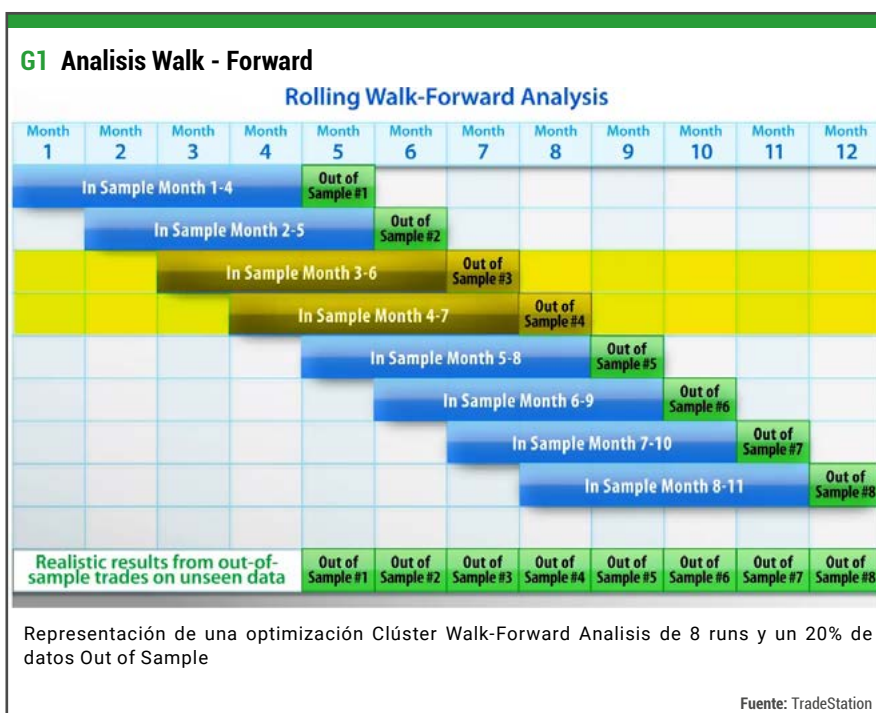
que nuestro sistema, tiene una ventaja sólida que explotar mediante la que ganar dinero operando en los mercados financieros. El método científico está presente en todo el proceso de desarrollo de sistemas y la optimización no es una excepción. Recordemos que el método científico exige que el objeto de estudio, nuestro sistema, sea observable objetivamente mediante un procedimiento ordenado y estandarizado, que nos permita evaluar el comportamiento de nuestro sistema aislando los distintos elementos del sistema. Recordemos que el único objetivo del trading es producir beneficios y que lo hace mediante la obtención de una ventaja que no es más que la cualidad o factor que te da superioridad frente al rival. Una gran ventaja debe ser:

- **Objetiva:** Lo contrario a subjetiva (no debería depender del trader que la opere)
- **Cuantificable:** Somos traders cuantitativos, todo tiene que basarse en datos.
- **Verificable:** La metodología, el protocolo nos lleva a un resultado, a una conclusión.
- **Consistente:** La ventaja no cambia, es la misma durante todo el procedimiento de evaluación y de operativa, es decir, mismas reglas para todos los trades.
- **Extensible:** Una buena ventaja es extrapolable a otros activos o mercados, las pautas universales y más sólidas funcionan en muchos mercados y/o timeframes.

Centrémonos ahora en el procedimiento de optimización.

Procedimientos de optimización. Walk-Forward clásico

La optimización se utiliza para mejorar una idea de trading, no para desarrollarla. La optimización walk forward clásica supone dividir al histórico en dos zonas, una zona más grande que usaremos en la optimización, que suele ser del orden del +/-70% (datos In Sample o INS) y otra más pequeña que



no se optimiza, en la que aplicaremos los parámetros que hemos obtenido en la optimización, que suele ser del orden del +/-30% (datos Out of Sample o OOS). Se recogen multitud de datos de rendimiento y riesgo de ambos segmentos. Uno de los más útiles es el Robustness index o Walk-Forward Efficiency que se calcula dividiendo el gradiente de la curva de rendimiento de la curva OOS entre el gradiente de la curva de rendimiento de la curva INS multiplicado por 100. En la práctica puede obtenerse dividiendo las tasas de retorno anualizadas de ambos periodos. Otros datos pueden ser el Draw Down u otros estadísticos como el profit factor, etc. Lógicamente, lo normal es que el periodo INS tenga una mayor significación estadística que el periodo OOS, pero este último debe tener un mínimo de trades. En una optimización Cluster Walk-Forward como la que veremos a continuación, donde disponemos de muchos periodos OOS (suelen ser al menos 10) se puede aceptar un número mínimo de 30 (si son 50 casi mejor), pero en esta fase donde solo disponemos de un periodo OOS, 30 me parece un número insuficiente. Menos de 150 me parece inaceptable y a partir de 300 en términos generales estaría bien. Esto dependerá mucho de los grados de libertad que tengamos en el sistema, no es lo mismo un sistema con un solo input que otro que tenga cinco. Tengamos en cuenta que, si tenemos 300 OOS y este supone el 30% del total, el perdido INS es de 700 trades. La comparación entre estos dos segmentos es un primer paso para evaluar la validez de la estrategia

Clúster Walk - Forward analysis

Este procedimiento lo relató Robert Pardo en 1992 en su libro "Design, Testing, and Optimization of Trading Systems". Lanzó una 2ª edición revisada en 2008: "The Evaluation and Optimization of Trading Strategies", una maravilla de

T1 Resultado de una optimización Cluster Walk-Forward Análisis

Cluster Analysis | Optimization Summary (In-Sample) | Walk-forward Summary (Out-Of-Sample) | Test Results | Graphs | IVL History | Performance Summary | Equity Preview | Distribution Analysis | Monte Carlo | Input Ranges

Highest Cluster Average: PASS** Centre: OOS% = 10 Runs = 6 Matrix Average: PASS
 Optimal Re-optimization Interval: 455 days Display: Walk-forward Overall result Tests per WFA: 2954
 320 bars Type: Rolling WFA with % based Out-Of-Sample 3D view

OOS% \ Runs	5	6	7	8	9	10	11	12	13	14	15	16	17
5	PASS **	PASS **	PASS **	PASS **	PASS **	PASS **	PASS **	PASS **	PASS **	PASS **	PASS **	PASS **	PASS
10	PASS **	PASS **	PASS **	PASS **	PASS **	PASS **	PASS **	PASS	PASS	PASS **	PASS	PASS	PASS **
15	PASS **	PASS **	PASS **	PASS **	PASS **	PASS **	PASS **	PASS	PASS **	PASS	PASS	PASS	PASS
20	FAILED	FAILED	PASS **	PASS	PASS **	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS
25	FAILED	FAILED	FAILED	PASS **	PASS **	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS
30	FAILED	FAILED	PASS	PASS	FAILED	PASS	PASS	PASS	PASS	PASS	PASS	PASS	PASS

Podemos ver el resultado resumen de una optimización WFA real en un sistema nuestro tendencial. En este caso elegimos un rango de 5 a 17 runs en incrementos de 1 en 1 y de porcentaje OOS entre 5% y 30% con incrementos de 5 en 5.

Fuente: TradeStation y elaboración propia

T2 Walk Forward Runs

	Inicio	Fin	Incremento
Out-Of-Sample %	10	30	5
Walk Forward Runs	5	30	5

Se observa claramente que hay una estacionalidad alcista, línea negra, desde octubre a finales de año.

Fuente: <https://spreadregreycarr.carrd.co/>

run (en total hay 8 runs) y cada uno de esos runs está formado por un 80% de datos INS y un 20% de datos OOS. La última fila de la imagen muestra cómo vamos recogiendo los datos OOS que se van generando, de tal forma que cubrimos una gran parte del histórico. Solo los primeros 4 meses quedan como INS. Los 8 restantes tienen datos OOS. Insisto en que esto es solo 1 clúster. Para hacer este tipo de optimización WFA debemos elegir un rango de runs y de porcentaje OOS en que realizaremos los tests para así generar muchos clústeres, cada uno con su configuración de runs y % OOS. En la tabla 2 vemos cómo elegir este rango que en esta tabla recogería 30 posibilidades o clústeres en total.

La mayoría de los clústeres tienen un PASS, lo que indica que han pasado las pruebas. Veamos cuáles son las pruebas que debe cumplir cada clúster:

- La suma de todos los runs OOS es positiva, es decir, ese clúster gana dinero en OOS.
- La media del Walk-Forward Efficiency de todos los runs OOS es del 50% o mayor.
- El 50% o más de todos los runs OOS fueron rentables.
- Muestra una distribución uniforme de los beneficios, es decir, ningún run ha aportado el 50% o más del beneficio neto total.
- Ningún run tiene un Draw down superior al 40%.

Para ser PASS debe pasar TODAS estas pruebas. Solo que una no se cumple, el clúster queda como FAILED. Para que la prueba sea PASS en su conjunto tienen que haber una mayoría de clúster PASS. Esta prueba la hacemos con una aplicación llamada Walk-Forward Optimizer (WFO) que forma parte de la plataforma TradeStation, aunque hay otras herramientas similares. También puede hacerse con Excel manualmente, aunque es un proceso realmente complejo. La herramienta nos permite recoger y ver mucha más información, incluso ver la curva de performance de todos estos datos que recoge un primer tramo INS y los runs de OOS que correspondan, como puede verse en la imagen 2.

Brac

“Tengo una manera relativamente fácil y robusta para que

libro que sigue totalmente en vigor. Como hemos visto, la optimización Walk-Forward es un proceso de comparación del rendimiento de la estrategia entre dos conjuntos de datos, uno que es optimizado por el optimizador (INS) y otro que no se optimiza (OOS). Durante un Análisis Clúster Walk-Forward (WFA), se realizan múltiples ejecuciones walk-forward. Para cada ejecución o run del Walk-Forward, el rendimiento de los datos Out-Of-Sample se comparará con el rendimiento de los datos In-Sample. El optimizador de walk-forward aplicará diferentes pruebas a los datos OOS de cada “run” para evaluar la consistencia del sistema y, basándose en este análisis, proporcionará una calificación global de la robustez del sistema. La potencia del Clúster Análisis radica en que mediante la elección de distintos números de runs y distintos porcentajes OOS, podemos comprobar cómo hubiera ido el sistema en muchos cortes distintos del histórico. La acumulación de estos datos es la fortaleza de esta prueba ya que estresa al sistema al máximo al evaluarlo en ventanas distintas de periodos INS/OOS. En la tabla 1 podemos ver el diagrama general de optimización WFA de uno de estos clústeres. Recoge un periodo de 12 meses, pero esto es solo un ejemplo, el periodo podría ser perfectamente otro. Toda la imagen recoge un solo clúster, es decir una parte de todo el proceso. En concreto recoge una optimización de 8 runs con un 20% de datos OOS. Cada fila de la imagen representa un

Para que la prueba sea PASS en su conjunto tienen que haber una mayoría de clúster PASS

se convenza de si un sistema que está desarrollando es mínimamente ajustado a la curva, o no. Se llama "Build, ReBuild and Compare" o BRAC. El método implica el uso de todos los datos históricos de su muestra de desarrollo para el instrumento (s) que tiene la intención de operar." Con esta contundente frase Keith Fitch en 2013 en su libro "Building reliable trading systems" su método BRAC. Fitch es el autor de Aberration en 1986, un sistema clásico entre los clásicos que estuvo muchos años considerado como uno de los mejores sistemas. Fitch critica con dureza

prácticamente muchas ideas clásicas sobre backtest y algunas sobre trading:

- Afirma que el Walk-Forward es poco menos que una secta sin ninguna base científica.
- Afirma que Monte Carlo está sobrevalorado como indicador de cuanto realmente ganará una estrategia o cual será el peor Draw Down.
- Critica también las divergencias técnicas como señales de entrada en el mercado.
- Incluso habla de una conspiración de las empresas que fabrican software de trading.

Él sostenía que no hacen falta análisis fuera de muestra ni complejos protocolos como el de Pardo, tan solo su método BRAC:

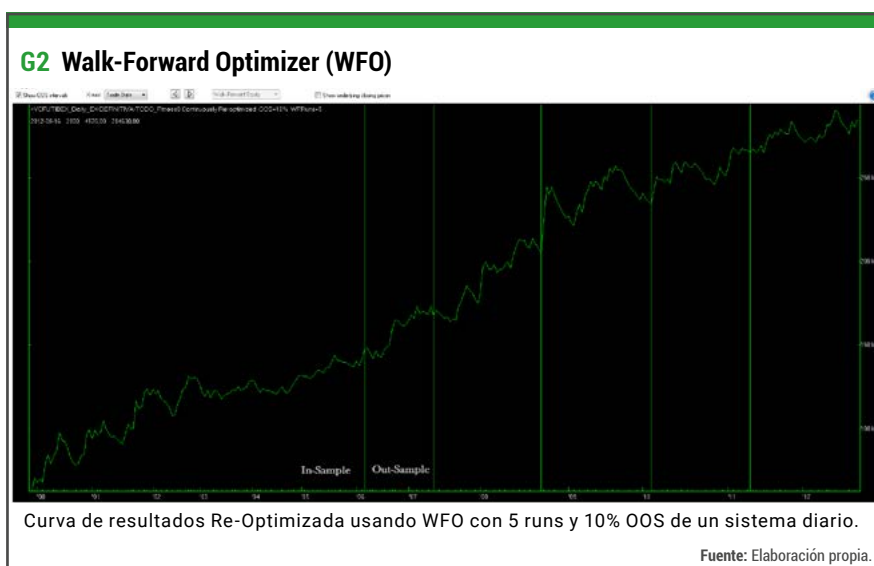
Construir (build)

A medida que construya la estrategia, lleve un registro de los pasos que dio y de la lógica de selección que utilizó, y si termina desarrollando una estrategia que tenga un buen rendimiento.

El conocimiento del sistema es clave para su confianza y para que no se ajuste a la curva. Ideas sencillas, lógicas y universales que funcionen en muchos activos. Para hacer el backtest o optimización usaba toda la muestra. No dividía el histórico en datos INS y OOS.

Reconstruir (rebuild)

A continuación, recorte algunos datos del final de los datos históricos. Reconstruya su estrategia utilizando los mismos pasos y criterios de decisión que utilizó la primera vez. Es decir, use los mismos rangos de optimización, las mismas restricciones y funciones objetivo.



Todo exactamente igual excepto eliminar una pequeña parte de los datos históricos. Por ejemplo, si está utilizando 20 años de datos de barras diarias, puede cortar el último año.

Comparar (compare)

Compara el rendimiento durante el período de los datos retenidos con el rendimiento de la primera estrategia durante el mismo período. Si el rendimiento es similar, puede tener bastante confianza en que su estrategia se mantendrá en la operativa real.

Conclusiones

En mi opinión Fitch usa en realidad un método clásico al que le añade un punto especial que tiene sentido para diferenciarse. Una muestra de 4000 trades de una idea sencilla, universal, que funciona en muchos activos, con pocas reglas y pocos inputs, es probablemente robusta optimices cómo optimices.

Fitch presenta en su libro Backtests de hasta 11.000 trades y además los backtestea en varios activos a la vez. En este contexto, si obtienes buenas curvas con riesgo controlado, es muy probable que el sistema sea robusto. Cumple dos criterios básicos para considerar que una muestra es válida. Significación estadística por tener muchísimos trades y representatividad del universo objeto de estudio ya que usa históricos de muchos años y encima mezclando activos. El método KISS, Keep It Simple Stupid, es básico en el desarrollo de sistemas. De hecho, Pardo también recomienda reglas sencillas y comprensibles.

También el libro de Fitch es muy interesante. BRAC es una pequeña parte de él, ya que desarrolla algunos sistemas e ideas y también la gestión monetaria.