



Universidad de San Andrés

Escuela de Administración y Negocios

Magister en Finanzas

Nombre del Trabajo Final de Graduación:

Arbitraje estadístico en terna de activos con redes neuronales

Autor: Procopio, César Horacio

D.N.I: 24.935.534

Director del Trabajo Final de Graduación: Basaluzzo, Gabriel Arturo

C.A.B.A. - Julio de 2019

Índice

1. Introducción	1
1.1. Planteo del problema	1
1.2. Marco teórico	2
1.3. Motivación	5
1.4. Organización del trabajo	7
2. Antecedentes	7
2.1. Arbitraje estadístico	7
2.2. Metodologías empleadas	8
2.2.1. Métodos basados en distancias	10
2.2.2. Métodos basados en cointegración	10
2.2.3. Métodos estocástico	11
2.2.4. Casos de aplicación	11
2.2.5. Otros posibles casos de aplicación	11
3. Metodología	13
3.1. Datos empleados	13
3.1.1. Datos simulados	13
3.1.2. Datos reales	15
3.2. Algoritmo	16
3.2.1. Objetivo de la red	16
3.2.2. Características y estructura de la red	17
3.2.3. Como se ajustan los pesos y los ponderadores. Etapa de entrenamiento	17
3.2.4. Función de error y criterios de convergencia	18
3.2.5. Ventanas temporales para entrenar y testear	18
3.3. Etapa de testeo y ejecución	19
3.3.1. Análisis de los errores y umbrales de ejecución	19
3.3.2. Presentación de nuevas observaciones	20
3.3.3. Ejecución de la estrategia	21
3.4. Como se obtienen las ganancias	21
3.5. Resumen de los parámetros del método	22

4. Resultados	23
4.1. Resultados con datos simulados	23
4.2. Resultados con datos reales	27
4.3. Ternas simuladas con datos reales	27
4.4. Resultados con la serie completa de precios observados	29
5. Conclusiones	31
6. Anexo	33
6.1. Estructura gráfica de la red	33
6.2. Estructura matricial de la red	34
6.3. Umbrales de equilibrio	36
6.4. Ternas de precios simuladas con reversión	37
6.5. Ternas de precios simuladas sin reversión	38
6.6. Ternas de precios simuladas con datos reales	39
6.7. Tabla de resultados con datos simulados con reversión	40
6.8. Tabla de resultados con datos simulados sin reversión	41
6.9. Tabla de resultados con datos simulados de bancos	42
7. Bibliografía	43



Lista de figuras

1.	Distribución de los excesos de retorno con reversión	25
2.	Distribución de los excesos de retorno sin reversión	25
3.	Dispersión de los excesos de retorno con reversión	26
4.	Dispersión de los excesos de retorno sin reversión	26
5.	Distribución de retornos para el caso real	28
6.	Dispersión de retornos para el caso real	28
7.	Estructura de la red neuronal	33
8.	Umbrales y zonas de equilibrio	36
9.	Simulación con reversión 1	37
10.	Simulación con reversión 2	37
11.	Simulación con reversión 3	37
12.	Simulación con reversión 4	37
13.	Simulación sin reversión 1	38
14.	Simulación sin reversión 2	38
15.	Simulación sin reversión 3	38
16.	Simulación sin reversión 4	38
17.	Simulación de bancos 1	39
18.	Simulación de bancos 2	39
19.	Simulación de bancos 3	39
20.	Simulación de bancos 4	39

Lista de cuadros

1.	Comparación de resultados en pruebas entre datos simulados con y sin reversión	23
2.	Excesos de retorno promedio en las simulaciones	24
3.	Resultados en pruebas con datos reales	27
4.	Excesos de retorno promedio en el caso real	27
5.	Resultados en pruebas con serie original	29
6.	Excesos de retorno promedio con serie original	29
7.	Resultados en pruebas con series original	30
8.	Simulaciones con reversión	40
9.	Simulaciones sin reversión	41
10.	Simulaciones con precios reales	42



Universidad de
San Andrés

Agradecimientos

Agradezco a la Universidad de San Andrés por impulsar espacios de formación y a todas las personas que participan para que ello sea posible. Quiero dar las gracias especialmente a mi tutor, Gabriel, quién no sólo me brindó toda la confianza y acompañó en esta ardua labor de realizar una tesis, sino que además fue un amigo y maestro en todo este proceso, y lo seguirá siendo. Debo también agradecer a mi familia por haberme motivado a iniciar una carrera universitaria y siempre me apoyaron en mis constantes estudios, especialmente a mi vieja a quien extraño mucho y en este momento estaría orgullosa de mi una vez más. No puedo dejar de mencionar a mis amigas y amigos de la vida quienes me apoyaron y brindaron su colaboración hasta las últimas palabras que escribí en este trabajo.



Universidad de
San Andrés

Resumen

Proponemos medir la posibilidad de realizar arbitraje estadístico con tres activos empleando una original estructura de red neuronal entrenada con tres series de precios. Las predicciones representan al valor de precios sintéticos producto de una combinación lineal entre un par de ellos que replican al otro en la terna. Las tres series de precios sintéticas resultantes representan un sistema de equilibrio. Encontrado el sistema de equilibrio, cada precio observado será comparado con el del precio sintético que lo representa. Para obtener ganancias proponemos vender un activo si este presenta un precio estadísticamente mayor al sintético. Con el monto vendido compramos una cartera formada por los otros dos activos. Por otro lado, si un activo presenta un precio menor a su esperado compraremos el mismo con lo obtenido de vender a los otros dos. La estrategia concluye cuando los tres activos revierten a su situación histórica de equilibrio, momento en el que los activos comprados presentarán mayor rendimiento que los vendidos. La hipótesis de la tesis es que se obtendrán sistemáticamente excesos de retornos sólo si las ternas revierten a un equilibrio. El objetivo principal de la tesis es sistematizar un método que mida el potencial que presenta una terna de precios para explotar oportunidades de arbitraje estadístico. Para determinar la validez del método, proponemos realizar diferentes pruebas con ternas de activos simulados con y sin reversión a un equilibrio. Los resultados evidencian que el método diferencia las características de cada grupo. Cuando observa datos que revierten a un equilibrio, detecta las oportunidades de ganancias. De esta forma, formalizamos una técnica para encontrar un sistema de equilibrio, emitir alarmas cuando en los precios se observen discrepancias y explotarlas a los fines de obtener mayores ganancias.

Palabras claves: Arbitraje estadístico, redes neuronales, terna de activos, excesos de retornos, reversión, sistema de equilibrio, rebalanceo de portafolio, ventas en descubierto.

1. Introducción

1.1. Planteo del problema

Comenzamos este trabajo preguntándonos si es posible medir las posibilidades de obtener excesos de retornos empleando arbitraje estadístico con tres activos utilizando redes neuronales. La red tiene la función de encontrar un sistema de equilibrio, detectar eventuales discrepancias en los precios y emitir alarmas cuando estas ocurran. Definimos exceso de retorno al diferencial que existe entre los rendimientos que obtenemos en una cartera que compramos comparados con los rendimientos de otra cartera que vendimos.

Para responder la pregunta que nos planteamos, vamos a poner a prueba el método que proponemos ante dos grupos de ternas de activos simuladas con características diferentes. Para un grupo (Grupo A), realizaremos simulaciones con un nivel de reversión a un sistema de equilibrio. Por el otro lado, simularemos ternas sin reversión (Grupo B), las que podrán moverse en forma más volátil y aisladas de un equilibrio. Posteriormente los resultados serán comparados para validar que en las condiciones límite de cada grupo el método funciona correctamente.

Hipótesis: La hipótesis de la tesis afirma que el comportamiento en ambos grupos será diferente. El método va a encontrar un sistema de equilibrio más o menos fuerte dependiendo de la relación en los precios y detectará las discrepancias si estas se presentaran. Observaremos excesos de retornos sistemáticos en las ternas que reversion a un equilibrio, mientras que en las otras obtendremos excesos de retornos aleatorios. Para que el método funcione la reversión debe estar presente, más que la correlación y la cointegración.

En el Grupo A, la red va a encontrar una relación de equilibrio fuerte y al presentarse discrepancias la red las detectará y emitirá una señal. Cuando los precios converjan obtendremos excesos de retornos en forma sistemática porque restablecido el equilibrio la cartera que compramos tiene que subir más que la que vendimos.

En el Grupo B esperamos obtener un sistema de equilibrio débil y volátil en el tiempo. Los activos no necesariamente convergen a un equilibrio y los excesos de retornos son aleatorios.

Objetivos: El objetivo principal de la tesis es sistematizar un método que mida el potencial que presenta una terna de precios para explotar oportunidades de arbitraje estadístico y obtener excesos de retorno. Este método también puede ser empleado para explotar las posibilidades en ternas que observamos buenas chances de arbitrar.

El método identifica ternas arbitrables y otorga la posibilidad de aprovechar las oportunidades. Las acciones concretas para lograr el objetivo principal consisten en aplicar el método creado en el Grupo A, (diseñado con reversión a un equilibrio) y registrar los excesos de retorno que se obtuvieron en él. De la misma forma, aplicamos el método en el Grupo B (diseñado sin reversión a un equilibrio). Finalmente, debemos analizar y diferenciar los resultados en ambos grupos y constatar que el método realizó las mediciones en concordancia con las condiciones de diseño.

Objetivos específicos: Simular datos para comparar en diferentes situaciones. Diseñar una red neuronal que encuentre un sistema de equilibrio y detecte las anomalías. Y definir una estrategia de ejecución cuando los precios rompan la relación de igualdad.

Si las pruebas realizadas verifican las hipótesis, se afirma la validez del método. El método diferencia ternas con buenas posibilidades de arbitrar (Grupo A) de las ternas que no las presentan (Grupo B). Demostraremos que el mismo identifica sistemas de equilibrio, emite las alarmas en el caso de detectar discrepancias y monitorea hasta la restauración de equilibrio. Las ganancias se obtienen cuando los precios convergen.

1.2. Marco teórico

Arbitraje estadístico es una estrategia desarrollada para obtener beneficios extras a partir de discrepancias que puedan existir en grupos de activos con respecto a una relación de equilibrio entre ellos. La detección de estas discrepancias se basa en la identificación de una combinación lineal de activos o la formación de un activo sintético cuya serie representa a otro activo. Este muestra un comportamiento que revierte al equilibrio con varianza finita. El activo sintético resume la relación de equilibrio entre un activo y el resto. Una discrepancia se manifiesta cuando el precio de un activo se diferencia significativamente de un precio esperado determinado por el activo sintético que lo representa.

La motivación principal del arbitraje estadístico, radica en que es posible predecir la trayectoria de los activos y de la existencia de una relación de equilibrio en el mercado. Esta

literatura, desafía a los principios en finanzas que dicen que los precios de los activos siguen caminos aleatorios e independientes, como dicen Lo y MacKinlay (1988)[1] o Guidolin et al (2009)[2]. Sin embargo, numerosos estudios han examinado las implicancias de la reversión a la media sobre el diseño de portafolios y administración de carteras de inversión. Podemos citar como ejemplo a Barberis (2002)[3] y Triantafyllopoulus y Montana (2009) [4].

Grupos de activos afectados por los mismos factores de riesgo presentan un comportamiento en su precio similar ya que cada factor de riesgo afecta de la misma manera a cada activo. Las diferencias que puedan surgir en su trayectoria pueden deberse a cambios estructurales (fusiones entre firmas, lanzamiento de nuevos productos, etc.), ineficiencias momentáneas del mercado, etc. Si los precios se comportan de manera semejante, entonces es posible encontrar una relación entre ellos explicada por la semejanza histórica. Cuanto más semejante es el movimiento de los precios, más fuerte es la relación que existe entre ellos. Un cambio estructural podría romper con el equilibrio y provocar una nueva relación entre los precios. Si este tipo de cambios se presente con más rapidez, las relaciones de equilibrios serán más volátiles. En la medida que este tipo de cambios no se observen, las relaciones de equilibrio serán más estables y más duraderas en el tiempo. Por otro lado, si el cambio se debe a ineficiencias del mercado, los precios deberían moverse en el corto plazo en dirección al equilibrio y su preservación. En este trabajo pretendemos desarrollar un método que encuentre una relación de equilibrio entre tres activos y aprovechar cuando se presenten discrepancias en las relaciones.

Un supuesto de este trabajo es que en una terna o grupo de activos que presente una relación de equilibrio en el largo plazo, tenderá a revertirse en el corto plazo ante cualquier desequilibrio en dicha relación y la misma se restablecerá nuevamente. Cuanto más fuerte es la relación, más rápida puede ser la reversión y por lo tanto mayor probabilidad de obtener éxito en la estrategia empleada.

En esta tesis proponemos emplear una red neuronal para encontrar la relación de equilibrio que existe en una cartera de tres activos. Cuando uno de los tres presente un valor discrepante con el equilibrio, la red emitirá una señal para ejecutar una estrategia de compra y venta para rebalancear la composición de la cartera. Los resultados medidos como retornos los obtenemos al comparar la cartera que compramos con la que vendimos. Realizaremos varias pruebas junto a sus respectivas comparaciones y de esta forma,

calcularemos las chances de éxito y los rendimientos esperados.

No buscamos predecir el precio de una acción, sino obtener ganancias con la reversión. Si la reversión se produce cuando el mercado esta en alza se habrán obtenido mayores ganancias. Si la reversión es cuando la tendencia del mercado esta en baja se habrán disminuido pérdidas.

La forma más tradicional de empleo de este tipo de estrategias es vendiendo en descubierto el activo que está por encima del equilibrio y comprar el activo que está debajo. En el mercado argentino un inversor puede tomar prestado una cantidad de un papel y venderlo en el mercado. A cambio, debe constituir una garantía (que puede ser dineraria, o bien un activo inmovilizado) y pagar una tasa anual (aproximada al tres por ciento) al propietario del activo alquilado. Ese préstamo tiene una duración determinada y puede renovarse hasta que quien vendió en corto considere que ya ganó lo suficiente con la baja y decida recomprar el papel para cerrar su posición. Este tipo de operaciones se reanudaron a partir de abril de 2018. Al momento de escribir esta tesis, dos acciones y tres bonos (Galicia, Pampa, AO20, AY24 y AA37) pueden operarse bajo esta modalidad. Además se puso un cupo máximo para el total de especies que todo el mercado puede vender en descubierto (Bustamante (2017) [5], Bezán (2018)[6]). Todo esto, muestra poco desarrollo en el mercado local para emplear el método que proponemos con ventas en descubierto. De todos modos, los resultados van a ser válidos para desarrollar este tipo de estrategias donde los mercados lo permitan.

Por otro lado, trabajar con datos de activos financieros de mercados emergentes o pocos desarrollados puede ser una tarea particularmente difícil. Existen en ellos cambios específicos que distorsionan los precios y las transacciones (informes inesperados, información privilegiada, poca liquidez, etc.). Sin embargo, decidimos trabajar con precios del mercado argentino (MERVAL), el cual en los últimos años ha tenido una volatilidad muy alta y una dependencia de los mercados más desarrollados. Estas características pueden presentar mayores discrepancias en los precios de los mercados emergentes que en mercados mas desarrollados y ultra arbitrados, brindando mejores oportunidades de obtener excesos de retornos explotando discrepancias entre los precios de grupos de activos.

1.3. Motivación

La tarea de hallar grupos de activos cointegrados no es fácil. En la práctica, son pocos los pares de activos que cumplen las características necesarias para emplear los métodos más conocidos. Entre otros motivos, esta ausencia da indicios que el mercado está muy arbitrado.

Creemos que también pueden existir otras series que si bien no están cointegradas, presentan un comportamiento cercano y con cierto nivel de reversión, brindando posibilidades de arbitrar. Habitualmente se suele trabajar con precios diarios y por lo tanto los test de cointegración o los métodos basados en distancias se realizan y miden en forma diaria. Si algunos activos presentan, por ejemplo, cointegración en precios semanales, los test tradicionales descartarían estas series cuando podrían ser potenciales para implementar una estrategia de arbitraje estadístico.

Por otro lado, cuando se trabaja con precios normalizados, estamos suponiendo que los precios están en una condición de equilibrio en el momento inicial de la serie. En este momento ambos precios se igualan a uno y la distancia es nula. Toda la estructura de distancia va a depender del momento en que la serie se inicie.

De esta manera, consideramos importante disponer de un método para testear a grupos de activos y medir rendimientos potenciales independientemente de las características observadas. Los métodos tradicionales detectan condiciones suficientes para emplear la estrategia pero no necesarias. Pueden existir otras situaciones donde la estrategia sea rentable. Desarrollar un test que detecte grupos de posible implementación de la estrategia parece una tarea no trivial. Consideramos valioso entonces el desarrollo de un método alternativo basado en pruebas empíricas e independiente de los test conocidos hasta el momento. En este trabajo proponemos un método para medir la probabilidad de ganancias en base a pruebas empíricas.

Predecir el comportamiento de los precios del mercado, o el de los activos que lo componen, es generalmente considerada una tarea compleja. Es bien conocida la hipótesis de los mercados eficientes (Fama (1970)[7]) la que sugiere que los precios reflejan inmediatamente toda la información disponible y que lo único que provoca cambios en los precios es la aparición de nueva información. Por lo tanto, como la incorporación de nueva información es impredecible, consideran que los precios en el mercado se comportan de manera total-

mente aleatoria. En consecuencia, parecería imposible obtener excesos de retorno a partir de la utilización de la información disponible. Acorde a esta hipótesis, la mejor decisión es no intentar predecir el comportamiento de los precios, simplemente consiste en comprar los activos y nunca venderlos. Sin embargo, la credibilidad que recibe esta hipótesis es un tema de debate para distintos investigadores como Haugen (1999)[8] y Los (2000)[9]).

Ante esta complejidad presentada para predecir el comportamiento de los precios de los activos sumada a la alta irregularidad y no linealidad que presentan las series financieras, los modelos lineales tradicionales, como por ejemplo ARIMA, no son considerados adecuados para abordar este problema. En efecto, los procesos ARIMA descansan en los supuestos que las series son estacionarias (o que pueden transformarse para cumplir con este supuesto) y que los errores son normalmente distribuidos (Box (1970)[10]). Infortunadamente, las series financieras no parecen cumplir con estos supuestos. Como resultado, existe una extensa literatura acerca de la efectividad de varias técnicas para predecir el comportamiento de los mercados. Como por ejemplo las Redes Neuronales, que muestran eficacia en la realización de un mapeo (dados los valores de entrada y de salida) y presentan la capacidad para aproximar cualquier función continua dado un nivel de precisión arbitrariamente definido (Cibenko (1989)[11]).

La motivación de este trabajo es sistematizar un método útil y original para medir los posibles excesos de retorno que podrían obtenerse dada una terna de activos. Los métodos tradicionales suelen emplearse para un par de activos y muy poca literatura se encuentra para grupos de más activos. Podemos encontrar abundantes métodos para arbitrar pares de activos y muchos de ellos de fácil implementación. Suponemos que en el caso de presentarse alguna oportunidad de arbitrar un par de activos, éstos no persisten mucho tiempo dado que sería fácil de arbitrar y restaurar el equilibrio.

También, es de suponer que a mayor cantidad de activos, mayores sean las posibilidades que uno de ellos salga de su equilibrio, en consecuencia mayores oportunidades de arbitrar se presentarían. Además, proponemos encontrar relaciones de arbitraje en ternas ya que un grupo de activos puede llegar a predecir mejor a otros. Por lo tanto, las relaciones de equilibrio serían más precisas. Incluso pueden encontrarse nuevas relaciones que en un par no se llegarían a observar.

Los métodos de conintegración son muy exigentes en cuanto a las condiciones. Ternas de activos pueden ser descartadas por estos test pero sin embargo, brindan buenas oportu-

nidades de arbitrar.

1.4. Organización del trabajo

El trabajo esta organizado de la siguiente manera: Continuaremos definiendo los principales aspectos y antecedentes del arbitraje estadístico. En la sección metodológica detallaremos como simulamos los datos que empleamos, realizaremos una breve descripción del funcionamiento de las redes neuronales y definimos como es ejecutada la estrategia. Posteriormente daremos detalle de las pruebas que realizaremos para validar el método. En la sección de resultados mostraremos gráficos y tablas mostrando los diferentes resultados que obtuvimos. Finalmente concluimos el trabajo haciendo una síntesis de lo realizado y las conclusiones a las que hemos arribado.

2. Antecedentes

2.1. Arbitraje estadístico

Según Avellaneda y Lee (2010)[12] el término arbitraje estadístico incluye estrategias de inversión que cumplen ciertas características. Estas son: Una regla sistemática para detectar señales de *trading*, la estrategia debe ser neutral al mercado en la esencia que no deben estar expuestas a riesgos de mercado, el mecanismo para obtener un exceso de retorno debe basarse en métodos estadísticos y la generación de señales debe estar basado en reglas. El éxito depende fuertemente de la modelización de las discrepancias o brecha entre los precios observados y los precios de equilibrio.

Muchos autores han sugerido diferentes métodos para obtener ganancias a partir de precios discrepantes tomando diferentes series de precios de activos financieros. Por ejemplo Burgess (1996)[13], Shadbolt y Taylor (2002)[14], Towers (2002)[15] trabajaron con precios del índice FTSE 100. Otros como Garrett (2001)[16], trabajaron sobre índices de contrato de futuros y Steurer y Hann (1996)[17] hizo lo mismo pero con tipos de cambio. Además, algunos de estos como [13], [14], [17] emplearon redes neuronales para modelar la dinámica de las discrepancias en los precios.

Diferentes estudios para pronosticar precios de activos han mostrado que la predicción en el precio de un activo puede mejorar significativamente si se modela una combinación de activos por sobre una serie individual (Shadbolt y Taylor (2002)[14]).

Reseña histórica El origen del arbitraje estadístico se acredita a un grupo de científicos informáticos, matemáticos y físicos reunidos por Morgan Stanley de Wall Street en el principio de los años ochenta. El equipo, que incluía a los científicos e informáticos se reunió para estudiar oportunidades de arbitraje en los mercados de valores, empleando modelos estadísticos avanzados y desarrollando un programa que opere en forma automatizada para explotar los desequilibrios del mercado.

Un aspecto fundamental de su investigación fue el desarrollo de métodos cuantitativos para identificar pares de valores cuyos precios exhibieran movimientos históricos similares, o que estuvieran altamente correlacionados. Si bien el modelo desarrollado se ejecutó con éxito en 1987, los siguientes dos años de operaciones tuvieron resultados lo suficientemente malos como para que en 1989 el grupo se disolviera.

A lo largo de los años, el método ha ganado atención entre los operadores individuales e institucionales y visto como una estrategia de inversión neutral al mercado. Esto se debe en gran parte al advenimiento de Internet y los avances en la tecnología comercial. Estos dos factores han ayudado a nivelar las reglas de juego para los inversionistas individuales, haciendo que los datos del mercado en tiempo real y las herramientas poderosas estén disponibles y sean accesible no solo para los comerciantes institucionales. Los participantes del mercado actual, ya sean comerciantes minoristas o un equipo de matemáticos altamente calificados, tienen acceso a datos del mercado financiero en tiempo real, sistemas de comercio de acceso directo, modelos avanzados de computadora y la capacidad de automatizar estrategias comerciales complejas.

Mediante el uso de la tecnología, así como los fundamentos, las probabilidades, las estadísticas y el análisis técnico, los operadores que emplean arbitraje estadístico intentan identificar las relaciones entre dos instrumentos, determinan la dirección de la relación y ejecutan las operaciones según los datos presentados.

2.2. Metodologías empleadas

Diferentes métodos se han propuesto para identificar pares de activos para aprovechar esta estrategia. La idea fundamental esta fundada en la existencia de una relación de equilibrio entre los activos. Si esta relación existe, entonces cualquier desviación que se presente debe ser revertida volviendo a su situación de equilibrio.

La forma más tradicional para identificar discrepancias en los precios de activos se hace

realizando una regresión en los precios de un activo a partir del precio de otros activos. Luego se prueba que los residuos de esa regresión presenten un comportamiento reversivo. Diferentes tipos de test han sido desarrollados para este propósito pero el más conocido es el de Dickey-Fuller and Phillip-Perron. (Enders (1995) [18]). Los residuos de la regresión realizada representan las discrepancias en cada momento de la serie en relación a los otros activos. El siguiente paso es crear un modelo que describa la dinámica de las discrepancias para medir las magnitudes de las mismas. Para obtener beneficios, el modelo debe estar acompañado de una estrategia. Una estrategia que emplea arbitraje estadístico, identifica puntos de entrada en los cuales las discrepancias son anormales y cuando estas se presentan debe tomar las acciones correspondientes. Una estrategia de arbitraje estadístico es considerada sin riesgo o también con ganancias en el largo plazo, pero esto depende fuertemente de la habilidad de los precios a retornar al estado de equilibrio histórico y sin discrepancias. Generalmente, cuanto mayor sea la presencia de discrepancias y mayor sea la velocidad con la que los activos revierten al equilibrio, mayores serán las ganancias de la estrategia.

Una de las estrategias más simples y conocida de arbitraje estadístico es llamada con el nombre de *Pairs Trading* la que se concentra sobre un par de activos. En este tipo de estrategia, no se está interesado en la tendencia de un activo en particular. El interés radica en la tendencia de la dinámica de la relación. La idea es que cuando dos activos se mueven en forma conjunta, es posible que existan algunas desviaciones en los movimientos individuales las cuales pueden ser oportunidades que se materializarán cuando la situación vuelva a su equilibrio. Esta estrategia consiste en detectar dos activos que se muevan históricamente en forma cercana esperando que alguno de ellos se desvíe. En este momento se recomienda comprar el activo que está por debajo del equilibrio y vender en descubierto el que está por encima. Los beneficios extras serán materializados cuando la relación vuelva al equilibrio.

Pairs Trading es entonces una estrategia diseñada para explotar ineficiencias en el mercado definidas como desviaciones en los precios de una relación de equilibrio observada en el pasado.

Existen principalmente tres métodos para emplear *Pairs Trading*: Método basado en distancias, método basado en cointegración y el método estocástico. Estos métodos varían básicamente en cuanto a la definición de la brecha entre los precios.

2.2.1. Métodos basados en distancias

El método basado en distancias fue propuesto por Gatev et al. (2006) [19] y fue empleado por varios autores, entre ellos Andrade et al. (2005) [20], Engelberg et al. (2009)[21], Do and Faff (2010)[22], Bowen et al. (2010)[23], y Broussard and Vaihekoski (2012)[24]. En este método, el movimiento conjunto es calculado a través de una medida de distancia, la cual es definida como la suma de los desvíos al cuadrado entre dos series de precios normalizadas. Una serie normalizada es definida con un valor igual a uno en su primer elemento, y evoluciona según los retornos originales.

Para desarrollar la estrategia se seleccionan dos activos con la menor distancia. Una posición en corto y largo es tomada cuando la distancia excede un límite determinado. Generalmente este límite es establecido como dos desvíos. Este umbral puede optar por otras alternativas, como por ejemplo tomar como parámetro una media móvil lo cual permitiría capturar alguna volatilidad temporal.

Desde un punto de vista práctico, el método es de fácil implementación e independiente de modelos econométricos, los que acarrearán una serie de problemas. Sin embargo, al ser una estrategia no paramétrica presenta una baja capacidad de predicción comparada con otros. El supuesto fundamental del método es que la diferencia de los activos presenta una reversión a la media. Toda divergencia indica un desequilibrio y la distancia es la medida de la misma.

2.2.2. Métodos basados en cointegración

El uso del concepto de cointegración se comenzó a usar por Lucas (1997) [25] y Alexander (1999)[26]. A partir de ellos se implementó en forma creciente y aplicado a la ingeniería financiera como por ejemplo Alexander y Dimitriu (2002)[27]; Bessler y Yang (2003)[28]; Yang et al. (2004)[29].

Cointegración es una técnica que permite modelar series temporales no estacionarias que presentan una tendencia estocástica en común. La ventaja que presenta el método de cointegración comparado con el de correlación, es que permite utilizar la información contenida en los niveles de las series. El supuesto fundamental de esta técnica es la reversión a la media de la diferencia de los precios de los activos. Sin embargo, en lugar de definir la diferencia entre los precios estandarizados de un conjunto de activos, la diferencia es definida con respecto al equilibrio a largo plazo de un sistema de precios cointegrados lo

que surge a partir de una combinación lineal entre ellos. Toda desviación del equilibrio debería revertir a la situación de equilibrio a largo plazo, implicando que una de las series se debe ajustar en orden de restablecer el equilibrio.

Usando técnicas de cointegración como teoría de base, la diferencia se genera a partir de los errores actuales con respecto a la situación de equilibrio a largo plazo. Si esos errores no presentan comportamiento estacionario y la relación de equilibrio existe, entonces estamos en presencia de un par cointegrado.

Los métodos más conocidos para testear la existencia de cointegración fueron propuestas por Engle and Granger (1987)[30] y Johansen (1988) [31]. El test de Johansen es un procedimiento para comprobar estadísticamente la existencia de series cointegradas. El método consiste en estimar una relación de cointegración de las series y testear la presencia de raíces unitarias en sus residuos con el test de Dickey y Fuller. La hipótesis nula de este test es la ausencia de cointegración en las series.

2.2.3. Métodos estocástico

Elliot et al (2005) [32], modelaron el proceso de reversión a la media de la brecha que existe entre dos activos, donde la brecha es definida como la diferencia entre los precios.

2.2.4. Casos de aplicación

En el trabajo de Caldas y Caldeira (2006)[33] proponen emplear precios de activos que cotizan en diferentes mercados (USA, Europa y Brasil), con la idea de comparar dos algoritmos para realizar arbitraje estadístico. Ellos eligieron tres mercados con características disímiles para poder realizar comparaciones de metodologías y testearlas bajo diferentes condiciones de mercado. Muchos trabajos existen sobre el tema, pero la diferencia que existen en los criterios de selección de activos, precios tomados en diferentes períodos hacen difícil la tarea de compararlas. Los autores concluyen que los rendimientos de las estrategias están influenciadas por las metodologías elegidas y que diferentes estrategias pueden ser útiles en diferentes mercados y diferentes períodos de tiempo.

2.2.5. Otros posibles casos de aplicación

- Rebalanceo de portafolios
- Seleccionar activos para formar una cartera

- Para arbitrar con ventas en descubierto
- Para eliminar riesgo de mercado
- Para arbitrar entre mercados
- ETF arbitrage



Universidad de
San Andrés

3. Metodología

3.1. Datos empleados

El método que proponemos vamos a llevarlo a la práctica con dos tipos de datos, emplearemos tanto datos simulados como datos reales.

Al primero de ellos lo emplearemos para comprobar que el algoritmo funciona correctamente y demostrar la hipótesis del trabajo. Con esto, buscamos demostrar que el modelo es capaz de encontrar una relación de equilibrio entre los precios, detecta irregularidades y cuando estas se presentan, emite las señales para ejecutar la estrategia.

Los precios reales los emplearemos para realizar un análisis exploratorio con este tipo de datos.

3.1.1. Datos simulados

El objetivo de la simulación de precios es comprobar que cuando se presentan series con reversión a un equilibrio, el método responde correctamente, y por lo tanto detecta las oportunidades de arbitrar. Por otro lado, cuando se le presenten ternas de precios sin reversión a un equilibrio, las oportunidades de arbitrar observadas por el método disminuyen.

Simulamos mediante el método de Montecarlo dos grupos de mil ternas cada uno. Las ternas que componen cada grupo presentan diferentes comportamiento.

En un grupo, que denominamos Grupo A, simulamos ternas que además de estar correlacionadas en los retornos, presentan grado de reversión a una ecuación de equilibrio. Para simular precios correlacionados empleamos la descomposición de Cholesky para generar números aleatorios. Empleamos una correlación para todos los pares de activos igual a 0.9.

Como ya mencionamos, los precios que componen esta terna deben presentar una reversión a un sistema de equilibrio que define las relaciones entre los activos. Dado que la red va a encontrar un activo sintético para cada activo de la terna, el sistema de equilibrio está formado por tres ecuaciones independientes. El mismo queda representado de la siguiente forma:

$$\widehat{S}_{1,t} = a_1 * S_{2,t} + b_1 * S_{3,t}$$

$$\widehat{S}_{2,t} = a_1 * S_{1,t} + b_2 * S_{3,t}$$

$$\widehat{S}_{3,t} = a_3 * S_{1,t} + b_3 * S_{2,t}$$

en forma general cada ecuación queda definida como:

$$\widehat{S}_{i,t} = a_i * S_{m,t} + b_i * S_{n,t}$$

donde:

$\widehat{S}_{i,t}$ es el precio de equilibrio del activo i en el momento t

$S_{i,t}$ es el precio observado del activo i en el momento t

$S_{m,t}$ y $S_{n,t}$ son los precios observados de los otros dos activos distintos de i

a_i y b_i son los coeficientes que multiplicados por S_m y S_n respectivamente, sintetizan a S_i

Para calcular el coeficiente a_i de cada ecuación del sistema generamos un número aleatorio entre 0 y $S_{i,1}/S_{m,1}$. A modo de ejemplo, si este número aleatorio es igual a cero, entonces el activo m no explica el comportamiento de i . Por el otro lado si el número aleatorio toma el valor máximo, el activo i es explicado totalmente por m .

$$a_i = \text{aleatorio}(\text{min} = 0, \text{max} = S_{i,1}/S_{m,1})$$

Para calcular el valor de b_1 se calcula por diferencia dado el valor del coeficiente a_i .

$$b_i = ((S_{i,1} - a_i * S_{m,1}) / S_{n,1})$$

para $i = 1, 2, 3$

Para cada terna simulada el sistema calculamos nuevamente el sistema de equilibrio. Hasta el momento, esto lo realizamos para simular datos y el método aún no participa.

La función estocástica que empleamos para simular la presentamos con la siguiente fórmula:

$$S_{i,t} = S_{i,t-1} * \exp[\mu - 0,5\delta_i^2 + \delta_i W_{i,t} - \alpha * (\log S_{i,t-1} - \log \widehat{S}_{i,t-1})]$$

donde:

S_t es el precio en t (aleatorio entre 1 y 3)

μ es la media de los retornos de i (0.002 para todos los activos simulados)

δ_i es el desvío de los retornos de i (0.004 para cada activo simulado)

$W_{i,t}$ es un proceso de Winner

α es la velocidad de reversión (0.02)

$\widehat{S}_{i,t-1}$ es el precio de equilibrio del activo i en $t - 1$

El término $\alpha * (\log S_{i,t-1} - \log \widehat{S}_{i,t-1})$ mide en que proporción se desviaron los precios observados $S_{i,t-1}$ con los precios de equilibrio $\widehat{S}_{i,t-1}$ en $t - 1$. Esa proporción multiplicada por el factor α representa el factor de reversión en t .

Los precios iniciales $S_{i,t=1}$ los determinamos en forma aleatoria con un número entre 1 y 3. Elegimos estos valores dado que las magnitudes no son relevantes y siendo números cercanos y no tan grandes la red aprende con mayor facilidad.

En el otro grupo, que denominamos Grupo B, simulamos precios sólo con la función estocástica ya que no presenta reversión al equilibrio. Los números aleatorios son generados con los mismos parámetros que en el Grupo A, guardando la salvedad que $\alpha = 0$.

La función estocástica queda definida de la siguiente forma:

$$S_{i,t} = S_{i,t-1} * \exp(\mu - 0,5\delta_i^2 + \delta_i W_{i,t})$$

Cada una de estas simulaciones las empleamos como insumo para el modelo para comprobar su correcto funcionamiento.

Los parámetros que elegimos se corresponden con los retornos de los precios de Banco Galicia, BBVA y Banco Macro que cotizan en el Merval.

Las figuras 9 a 16 se muestran algunos gráficos a modo de ejemplo para visualizar algunas simulaciones de ambos grupos.

3.1.2. Datos reales

Utilizaremos también el método para medir las chances de obtener excesos de retornos con tres bancos que cotizan en el mercado de valores argentino. Para esto, descargamos los datos desde Yahoo Finance de la terna de acciones compuesta por Banco Galicia, BBVA y Banco Macro desde el 1 de enero de 2012 al 30 de abril de 2019 (1784 observaciones). Se eligieron estos activos por el sólo hecho de presentar mayor liquidez.

Con estos precios, simulamos diferentes ternas para analizar posibles trayectorias. Calculamos los retornos de cada serie y para cada simulación los mezclamos en forma aleatoria de modo de alterar su orden pero guardando su misma correlación en el tiempo. A partir de estos retornos desordenados, reconstruimos las series de precios obteniendo diferentes simulaciones de precios con estructura semejantes a los observados en el mercado.

Las figuras 17 a 20 muestran gráficos a modo de ejemplo para visualizar algunas ternas de precios.

3.2. Algoritmo

Para encontrar la relación de equilibrio entre los precios de los activos, usamos el método de regresión no paramétrica denominado red neuronal.

Elegimos emplear redes neuronales porque las mismas presentan mayor capacidad de predicción que otras regresiones paramétricas. También son capaces de encontrar relaciones no lineales y además pueden adaptarse a nuevas relaciones de equilibrio en cada momento t ya que continúan aprendiendo en cada observación.

El tipo de red neuronal que empleamos es denominada de múltiples capas (en inglés *Multi-layer feed-forward*). Es un método supervisado dado que la red conoce los elementos a predecir y ajusta sus pesos de tal manera que las salidas de la red se le asemejen a las entradas. El método de aprendizaje de la red es denominado propagación en reversa (*back-propagation* en inglés). Este tipo de red consiste en una serie de neuronas ordenadas en capas. La primera capa se llama capa de entrada, la última capa se llama capa de salida y entre ellas se encuentran las capas ocultas. Cada neurona de una capa es conectada con todas las neuronas de la capa siguiente. La conexión entre neuronas es caracterizada por los pesos, los que reflejan los grados de importancia de cada conexión de la red. En el proceso de aprendizaje, los pesos van adoptando valores tal que minimicen la diferencia cuadrática entre lo observado y lo estimado por la red.

3.2.1. Objetivo de la red

El objetivo de la red neuronal es construir tres activos sintéticos que representen al sistema de equilibrio. La red neuronal ajustará los pesos contenidos en sus capas ocultas para construir, a partir de la combinación de dos activos y su respectivas ponderaciones, a cada activo sintético. Cuando los tres activos son semejantes a sus respectivos sintéticos, significa que los precios no presentan discrepancias con la relación histórica entre sí. Cuando uno de los tres activos se aleja del activo que lo representa, la situación de equilibrio se rompe momentáneamente. Esta situación provoca que la red emita una señal para ejecutar la estrategia. Una vez que la estrategia se ejecutó, hay que esperar que los activos reviertan a su situación de equilibrio histórica, momento en el que las ganancias

se verán materializadas.

3.2.2. Características y estructura de la red

La red neuronal que diseñamos para construir los activos sintéticos presenta la capa de entrada idéntica a la capa de salida. El precio de tres activos será presentado como entrada para predecir el precio de ellos mismos en la salida de la red. Cada uno de los tres activos sintéticos será producto de una combinación únicamente de los otros dos. El gráfico que representa a la red puede observarse en la figura 7 del anexo. La red funciona como si fueran tres regresiones independientes (una para cada activo), pero está diseñada como un único modelo integrado.

La red neuronal está compuesta por una capa de entrada, tres capas ocultas y una capa de salida. La capa de entrada contiene las series de precios de los tres activos elegidos con su debido tratamiento. La primer capa oculta contiene ponderadores, los que definen que fracción de cada activo se utiliza para replicar al sintético. No está de más aclarar que la suma de los ponderadores debe ser igual a uno. las otras dos capas ocultas contienen los pesos que son entrenados para que las neuronas se organicen y aprendan a reconocer distintas características del espacio total de entrada. Estas capas están compuestas por nueve y seis neuronas respectivamente. Finalmente, la capa de salida que contiene los precios de los tres activos a predecir.

3.2.3. Como se ajustan los pesos y los ponderadores. Etapa de entrenamiento

La etapa de entrenamiento comienza con valores de ponderadores igual a 0.5 y los pesos en las capas ocultas con valores aleatorios comprendidos entre cero y uno. Estos ponderadores y pesos fueron optimizados mediante el método de propagación en reversa. Este método emplea un ciclo de propagación de dos fases. La primera fase aplica un patrón a la entrada de la red como estímulo, este se propaga desde la primera capa a través de las capas siguientes de la red hasta generar una salida. La señal de salida se compara con la deseada y se calcula el error para cada una.

En la segunda fase, los errores se propagan hacia atrás, partiendo de la capa de salida hacia todas las neuronas de la capa oculta que contribuyen directamente a la salida. Las neuronas de la capa oculta solo reciben una fracción de la señal total del error, basándose aproximadamente en la contribución relativa que haya aportado cada neurona a la salida

original. Este proceso se repite, capa por capa, hasta que todas las neuronas de la red hayan recibido una señal de error que describa su contribución relativa al error total.

Este proceso lo realizamos en forma iterativa. Cada iteración se denomina época (en inglés *epoch*) y en la misma presentamos todas las observaciones del conjunto de entrenamiento. En cada época la red ajusta sus ponderadores y pesos de tal forma de reducir el error. A medida que el proceso continúa los ponderadores y pesos convergen gradualmente a un óptimo local. En una época, cada patrón del conjunto de entrenamiento es presentado a la red en forma aleatoria. Con este patrón, la red genera una salida y el error obtenido es propagado en forma inversa por toda la red ajustando los ponderadores y los pesos. Una vez que los errores se hayan propagado y los ponderadores y pesos se hayan modificado para ajustarse a la capa de salida, los ponderadores y pesos que tienen que valer cero son anulados. También, si la suma de los ponderadores es mayor a uno, se lo transforma para que la suma sea igual a uno. Cada época finaliza cuando todos los patrones de entrenamiento fueron presentados. Este proceso se llama "Modo de presentación de patrones" (conocido en inglés como *pattern-mode*).

3.2.4. Función de error y criterios de convergencia

La función de error que busca minimizar la red está determinada por una matriz de 1 x 3 (un error por cada serie) y es representada por la ecuación:

$$error = (Y - \hat{Y})^T * (Y - \hat{Y})$$

donde Y e \hat{Y} son matrices de $n \times 3$.

La red itera mientras el valor de cada elemento de la función de error disminuya respecto a la iteración anterior y que el número de iteraciones sea menor a un valor establecido.

3.2.5. Ventanas temporales para entrenar y testear

Cada serie de precio esta formada por 500 cotizaciones. Los primeros 300 precios los usamos para que la red encuentre la relación de equilibrio (etapa de entrenamiento de la red)

Durante las restantes 200 observaciones es cuando evaluamos a la estrategia (etapa de testeo de la estrategia).

3.3. Etapa de testeo y ejecución

Una vez que la red encontró una relación de equilibrio, aceptada según los criterios de convergencia, continuamos con la etapa de testeo y ejecución de la estrategia. Esta etapa consiste en tres principales pasos. El primero consiste en definir los umbrales que determinan una ruptura de la relación de equilibrio. En un segundo paso, presentamos a la red los nuevos precios que pertenecen a la etapa de testeo. Finalmente, en caso de encontrar un desequilibrio, ejecutamos la estrategia y se le da seguimiento.

Tradicionalmente, toda estrategia de arbitraje estadístico se desarrolla vendiendo en descubierto un activo y comprar con el monto de esta operación a los otros. Dado que en mercado argentino están limitadas este tipo de operaciones, para la estrategia que proponemos es necesario contar con un portafolio que contenga la terna de activos analizada.

3.3.1. Análisis de los errores y umbrales de ejecución

El primer paso para poner en ejecución el método consiste en definir los umbrales que delimitan las zonas de equilibrio. Estos valores los determinamos realizando un análisis de los errores en las predicciones durante la etapa de entrenamiento. Definimos dos tipos de límites, los umbrales exteriores determinan los márgenes que indican una ruptura de equilibrio cuando estos se traspasan, mientras que los umbrales interiores delimitan la banda que representa la reversión al equilibrio.

En la figura 8 del anexo se muestra un gráfico donde podemos ver una representación de un activo sintético y sus umbrales.

Cálculo de los umbrales: Los desequilibrios se producen cuando el error observado en uno de los precios sobrepasa los umbrales exteriores (superior o inferior) determinados por valores de errores extremos. Cada sintético representado por la red posee sus propios umbrales.

Dado que el método contempla los gastos de transacción es necesario que exista una separación mínima (ganancia mínima) entre los umbrales interiores y los exteriores.

Si se llegaran a presentar ternas que se mueven en forma muy cercana, es de esperar que los percentiles de error también estarían muy cercanos. En este caso, los excesos de retornos serán inferiores a los gastos de transacción. Para evitar esto y asegurar rendimientos

positivos establecemos una brecha mínima, la cual es un valor que fuerza una separación entre umbrales.

Los umbrales interiores se calcula para cada activo con el percentil 30 (umbral interior inferior) y el percentil 70 (umbral interior superior) de cada uno de sus errores. Estos valores se emplean para delinear una banda que indica reversión al equilibrio en el caso que se haya producido un desajuste previo. Si los errores están entre estos valores entonces las relaciones entre los precios está en equilibrio implicando un cierre de la estrategia.

Definimos al umbral exterior superior como:

$$\text{maximo}((\text{umbral}_{\text{interiorSuperior}} + \text{brechaMinima}), \text{percentil90})$$

Mientras que el umbral exterior inferior como:

$$\text{minimo}((\text{umbral}_{\text{interiorInferior}} - \text{brechaMinima}), \text{percentil10})$$

3.3.2. Presentación de nuevas observaciones

Una vez que definimos los umbrales que representan el limite de inicio y final de la estrategia, se presentan a la red una a una las nuevas ternas de precios. Con estos nuevos precios, la red calcula las estimaciones de equilibrio de los mismos. Si el error que surge de estas estimaciones y los precios exceden alguno de los umbrales, entonces la red emite una señal de ejecución de la estrategia.

Para el primer día ($t = 301$), la red evalúa si los precios están en concordancia con la relación de equilibrio. Si alguno de los precios no esta dentro de los umbrales de equilibrio, entonces la red emite una señal para operar y se abre una posición. Si los precios están en concordancia entonces la red no emite señal de alarma y procedemos a analizar el día siguiente ($t = 302$), así sucesivamente hasta que detecte alguna discrepancia.

Con cada nueva observación, la red recalcula todos sus pesos, lo que le permite adaptarse constantemente a posibles nuevas relaciones de equilibrio. Con este ajuste la red puede aprender si se presentó algún cambio estructural en los precios. Es decir que la relación que aprendió la red es dinámica, y cambia día a día con las nuevas observaciones.

Mientras la posición está abierta, la red continúa con el mismo proceso de observación y recálculo buscando que los precios se encuentren en equilibrio nuevamente o hasta observar los precios en $t=500$ (momento definido como fin de la ventana temporal). Cuando

encuentra nuevamente que los precios están en equilibrio, la estrategia es cerrada y la red continuará analizando las relaciones para nuevas posibles oportunidades de arbitrar.

3.3.3. Ejecución de la estrategia

Como mencionamos anteriormente, cuando uno de los tres activos sobrepasa uno de sus umbrales de tolerancia de equilibrio, la red emite una señal para ejecutar la estrategia. La red también emitirá una señal para operar cuando un precio de la terna presente un valor inferior al observado en el mercado en el momento t con un error menor al $umbral_{inf}$.

Cuando el error en la predicción supera el $umbral_{sup}$ la estrategia a implementar consiste en vender el activo cuyo precio supera la barrera de equilibrio y con el dinero proveniente de esa venta comprar la cartera que lo replica. Las proporciones en las que se debe fraccionar el monto proveniente de la venta lo determinan los pesos de la primera capa de la red.

Por el contrario, si el error de una predicción de un activo es inferior al $umbral_{inf}$ la ejecución de la estrategia la desarrollamos comprando un monto determinado del activo. Los fondos provenientes para comprar este activo los generamos vendiendo una cartera en proporción a los pesos indicados en la primera capa de la red.

3.4. Como se obtienen las ganancias

La estrategia la ejecutamos cuando un activo presenta en el mercado un precio significativamente mayor (a) o menor (b) al que determina la relación de equilibrio. En ambos casos los beneficios se verán materializados cuando los precios observados converjan nuevamente a una situación de equilibrio. Para cuantificar los resultados se compara la rentabilidad que obtuvo la cartera que compramos con la cartera que vendimos durante el período que la estrategia se mantuvo abierta.

En el caso (a) si la convergencia se produce con movimientos de precios ascendentes, entonces la cartera comprada tendrá mayor rendimiento que la vendida, dado que el activo que vendimos tendrá una rentabilidad menor a los dos que compramos.

En el caso (b), si la convergencia se produce con movimientos de precios descendentes, entonces la cartera comprada tendrá mayor rendimiento que la vendida, dado que el activo que compramos tendrá una rentabilidad mayor a los dos que vendimos. Si por el contrario

la convergencia se presenta con movimientos de precios descendentes, el par de activos que vendimos presentará una caída mas abrupta que el que compramos, minimizando perdidas.

3.5. Resumen de los parámetros del método

Parámetros de la red:

Cantidad de neuronas en las capas ocultas (9 en en la primera capa, 6 en la segunda)

Tasa de aprendizaje (0.05)

Ventana temporal de entrenamiento (300 días)

Ventana temporal de testeo (200 días)

Cantidad de iteraciones en entrenamiento (500)

Cantidad de iteraciones en ventana de testeo (50)

Función de activación (lineales)

Parámetros de la simulación de precios:

Correlación entre activos (0.9)

Velocidad de reversión (0.02)

Precios iniciales (aleatorio entre 1 y 3)

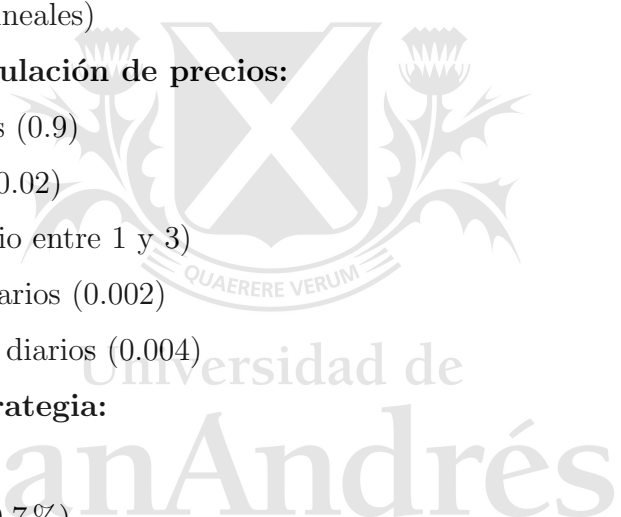
Media de los retornos diarios (0.002)

Varianza de los retornos diarios (0.004)

Parámetros de la estrategia:

Brecha mínima (0.05)

Gastos de transacción (0.7%)



4. Resultados

En este apartado vamos a mostrar los resultados que obtuvimos en las simulaciones que realizamos. Para demostrar la veracidad del método, realizamos pruebas con datos simulados como así también con datos reales de activos pertenecientes a tres bancos que cotizan en el Merval. Realizamos mil simulaciones con ternas simuladas en ambos grupos (A y B) y otras mil para ternas simuladas a partir de los retornos observados de los bancos.

4.1. Resultados con datos simulados

En una terna particular, el método puede no encontrar discrepancias en los precios dado que en la ventana de ejecución de la estrategia no se observaron desviaciones con respecto al equilibrio. En este caso el método no emite alarmas.

En los casos donde detecta discrepancias, el método emite una alarma para abrir la estrategia. Cuando los precios vuelven al equilibrio decimos que la estrategia se cerró, cuando los precios no convergen en la ventana temporal establecida (doscientos días), la estrategia finaliza y queda abierta. Cuando la estrategia se cierra los excesos de retornos deben ser positivos, dado que al converger, la cartera comprada presenta mayores retornos que la venta. Cuando la estrategia queda abierta los excesos de retornos pueden ser negativos o positivos. En el agregado total son 505 los casos con ganancias y 494 los casos con pérdidas.

En el cuadro 1 contabilizamos el comportamiento del método en las mil corridas que realizamos con los precios simulados con y sin reversión.

	Con reversión				∴	Sin reversión			
	Exc.Ret+	Exc.Ret.-	Total	∴		Exc.Ret.+	Exc.Ret.-	Total	∴
Cerradas	916	0	0	916	∴	266	0	0	266
Abiertas	57	26	0	83	∴	239	494	0	733
Sin discrepancia	0	0	0	0	∴	0	0	1	1
Total	973	26	0	1000	∴	505	494	0	1000

Cuadro 1: Comparación de resultados en pruebas entre datos simulados con y sin reversión

Como podemos observar, en la parte del cuadro correspondiente a los grupos de ternas con reversión, en la mayoría de los casos la estrategia queda cerrada y con excesos de retornos

positivos (916 casos). No se observan casos donde la estrategia cierre con pérdidas. En menor medida, el método detecta discrepancias pero como los precios no retornan al equilibrio la estrategia queda abierta (83 casos). En estos se observan tanto ganancias (57) como pérdidas (26).

Por otro lado, en la parte del cuadro correspondiente a las ternas sin reversión, observamos un comportamiento diferente. Los casos donde se rompió el equilibrio y la estrategia cerró con excesos de retorno positivo fueron 266. En 733 ocasiones los precios no volvieron al equilibrio y fueron más las veces que se obtuvo pérdidas (494 veces) que las que se obtuvo ganancias (239 veces). Se observa un caso donde los precios no presentaron discrepancia.

En el cuadro 2 mostramos los excesos de retorno promedio para cada caso para ambos grupos de activos.

	Con reversión			Sin reversión		
	Exc.Ret.+	Exc.Ret.-	Total	Exc.Ret.+	Exc.Ret.-	Total
Cerradas	7.7 %	0	7.7 %	7.1 %	0	7.1 %
Abiertas	8.9 %	-4.3 %	4.7 %	10.7 %	-10.2 %	-3.3 %
Total	7.8 %	-4.3 %	7.5 %	8.8 %	-10.2 %	-0.5 %

Cuadro 2: Excesos de retorno promedio en las simulaciones

Observamos que en el grupo con reversión, las estrategias cerradas arrojaron un retorno promedio de 7.7 %. Incluso en las ocasiones donde la estrategia quedó abierta el retorno promedio total fue positivo (4.7 %). Observamos además que la media total de las mil corridas es igual a -0.5 %, valor muy cercano al 0.7 % definido como gasto de transacción. Este valor que puede significar la indiferencia de poder emplear el método en este grupo.

En el grupo sin reversión, las estrategias que no cerraron presentan más amplitud en la distribución de los excesos de retorno arrojando valores promedios que van desde el 10 % al -10 %.

En la figura 1 mostramos dos histogramas superpuestos con la distribución de los excesos de retornos en las ternas que presentan reversión. Un histograma representa a los casos en que la estrategia cerro y el otro a los que la estrategia quedó abierta. La figura 2 hace lo mismo para el grupo sin reversión.

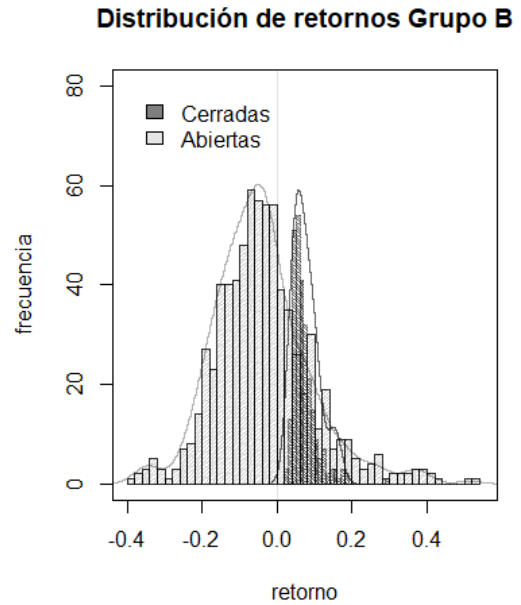
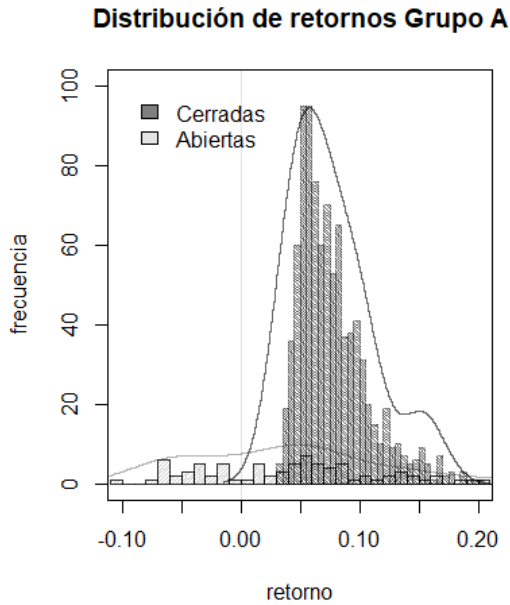


Figura 1: Distribución de los excesos de retorno con reversión

Figura 2: Distribución de los excesos de retorno sin reversión

En el Grupo A podemos ver mayor frecuencia de estrategias cerradas, todas distribuidas con ganancias mayores a cero. En cambio, en el Grupo B observamos mayor ocurrencias de casos abiertos y los retornos se distribuyen alrededor del cero con mayor dispersión.

En la figura 3 y 4 mostramos la relación entre los excesos de retornos y el tiempo transcurrido hasta que la estrategia se cerrara. Estas figuras nos brindan una noción del tiempo que tarda una terna en converger.

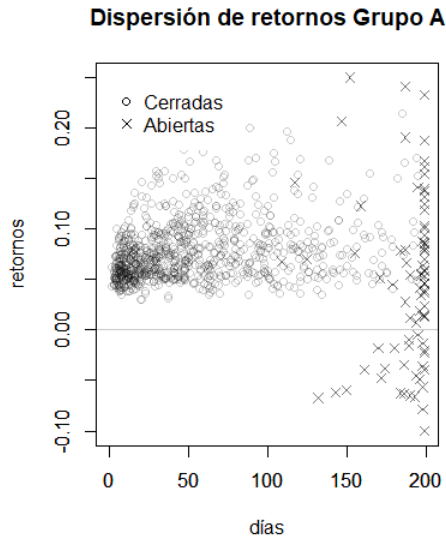


Figura 3: Dispersión de los excesos de retorno con reversión

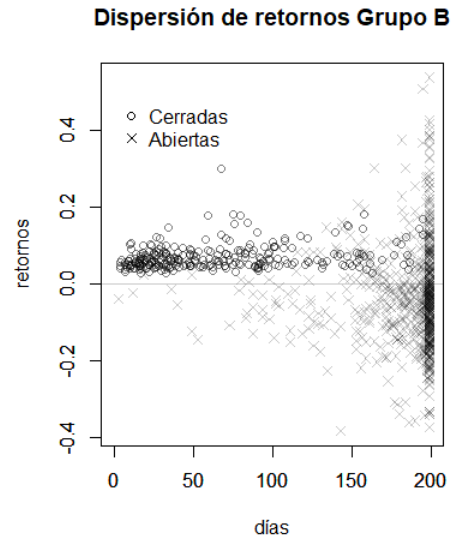


Figura 4: Dispersión de los excesos de retorno sin reversión

No observamos un patrón que indiquen una relación entre el retorno y el tiempo transcurrido. En el Grupo A observamos una concentración de puntos en los primeros 50 días. Esto está en línea con el factor de reversión igual a 0.02. En el Grupo B predominan las estrategias abiertas y los excesos de retorno positivos se compensan con los negativos.

En el anexo los cuadros 8 y 9 mostramos cuarenta de las mil simulaciones correspondientes al Grupo A y Grupo B respectivamente. En la misma podemos observar los siguientes campos:

Compra: Indica el valor final de la cartera comprada (por cada unidad de dinero comprado)

Venta: Indica el valor final de la cartera vendida (por cada unidad de dinero vendido)

Inicio: Indica el momento en que se ejecutó la estrategia

Final: Indica el momento en el que se cerró la estrategia. Si es 500 indica que la estrategia permaneció abierta finalizada la ventana de testeo

t: Indica el tiempo que la estrategia permaneció abierta

Ex.Ret.: Indica el exceso de retorno de la estrategia

Ex.Ret. Anual: Indica el exceso de retorno anualizado

Estado: Indica el estado final de la estrategia (abierta o cerrada)

Ganancia: Indica si la cartera rindió positivamente

p-value: Es el valor del test de Dickey-Fuller para testear si la terna está cointegrada

4.2. Resultados con datos reales

4.3. Ternas simuladas con datos reales

A modo de ejemplo vamos a presentar algunos resultados simulando ternas de precios a partir de los retornos observados. Una terna de retornos es seleccionada al azar formando una serie de retornos con la que se reconstruye la serie de precios.

En el cuadro 3 contabilizamos los resultados que surgieron al aplicar el método a los precios de los bancos. Si bien los casos donde la estrategia quedó abierta (624 veces) supera a las que cerraron (376 ocasiones), la cantidad de casos que presentan excesos de retorno positivo alcanzan al 65.7% del total.

	Exc.Ret.+	Exc.Ret.-		Total
Cerradas	375	1	0	376
Abiertas	282	342	0	624
Sin discrepancias	0	0	0	0
Total	657	343	0	1000

Cuadro 3: Resultados en pruebas con datos reales

Al observar los excesos de retornos promedio para cada caso en el cuadro 4, observamos que en el total general se obtiene un 4.4% de ganancias. En los casos donde la estrategia permanece abierta, en promedio los excesos de retornos positivos (16.4%) son superiores a los negativos (-13.5%).

	Exc.Ret.+	Exc.Ret.-	Total
Cerrada	11.9%	-0.01%	11.8%
Abierta	16.4%	-13.5%	0.01%
Total	13.8%	-13.4%	4.4%

Cuadro 4: Excesos de retorno promedio en el caso real

En la figura 5 mostramos la distribución de los excesos de retornos donde podemos notar el sesgo a los valores positivos. Análisis similar podemos realizar en la figura 6. Vemos en ambas que la distribución de retornos en las estrategias que cerraron están concentradas alrededor del 12% mientras que las que permanecen abiertas presentan mayor dispersión

y aleatoriedad.

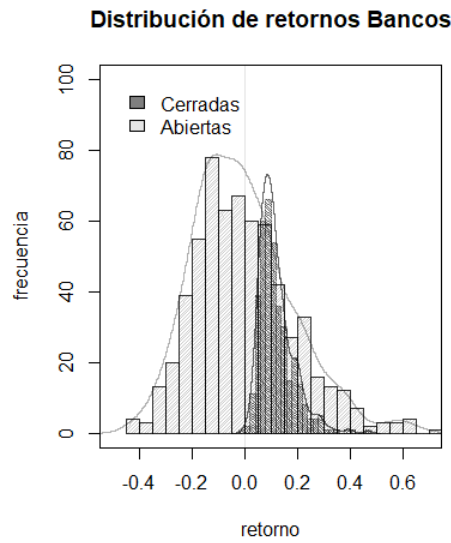


Figura 5: Distribución de retornos para el caso real

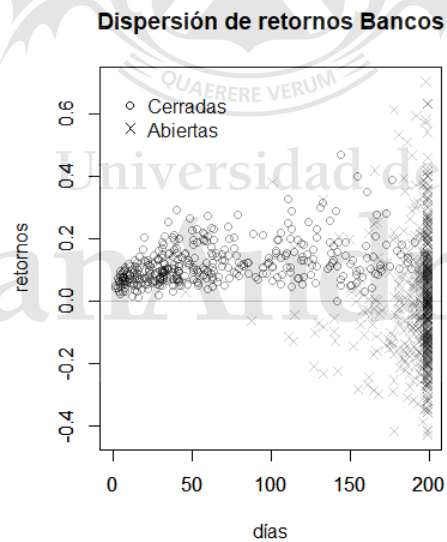


Figura 6: Dispersión de retornos para el caso real

En el cuadro 10 del anexo mostramos cuarenta de las mil simulaciones correspondientes a las pruebas realizadas con bancos.

4.4. Resultados con la serie completa de precios observados

Como hemos visto, la terna de bancos analizados no presenta muy buenas chances para ejecutar la estrategia. A continuación vamos a mostrar los resultados que obtuvimos con la serie de precios observadas, sin modificación. Estos resultados nos van a decir cuanto hubiéramos ganado si implementáramos la estrategia sobre esta terna. Comenzamos empleando los primeros 300 días para entrenar a la red y los siguientes 100 días para testear el equilibrio encontrado. Cuando el método emite una alarma se ejecuta la operación hasta que las ternas vuelvan al equilibrio o finalice la ventana establecida. Cuando una de estas dos opciones se presente (cierre de estrategia), continuamos analizando el equilibrio esperando nuevas alarmas. Este proceso continúa hasta finalizar la serie original.

En el cuadro 5 observamos que el método ejecutó diecinueve alarmas. En nueve de ellas las ternas revertieron al equilibrio. En las diez restantes la estrategia quedo abierta y los precios no se equilibraron. Si bien las observaciones parecen mejorar a la simulación anterior, la terna no muestra evidencia para suponer que es buena candidata.

textbf	Exc.Ret.+	Exc.Ret.-	Total
Cerradas	9	0	9
Abiertas	4	6	10
Total	13	6	19

Cuadro 5: Resultados en pruebas con serie original

En el cuadro 6 mostramos los excesos de retorno promedio para cada caso. Las diferencias en las ganancias entre pérdidas y beneficios no parecen significativas.

	Exc.Ret.+	Exc.Ret.-	Total
Cerradas	4.6 %	0 %	4.6 %
Abiertas	8.8 %	-7.8 %	-11.4 %
Total	5.9 %	-7.8 %	1.5 %

Cuadro 6: Excesos de retorno promedio con serie original

En el cuadro 7 mostramos el desarrollo de la ejecución de las estrategias detallando el rendimiento en cada una de las diecinueve alarmas.

Alarma N°	Inicio	Cierre	t	Estado	Comprada	Vendida	Ex.Ret.	Ganancia
1	359	364	5	Cerrada	1.082	1.023	5.7%	Gana
2	381	436	55	Cerrada	2.257	2.141	5.4%	Gana
3	440	442	2	Cerrada	0.972	0.927	4.8%	Gana
4	451	543	92	Abierta	1.161	1.258	-7.7%	Pierde
5	544	645	100	Abierta	1.568	1.281	22.4%	Gana
6	646	736	90	Cerrada	1.239	1.222	1.3%	Gana
7	738	771	33	Cerrada	1.239	1.177	5.2%	Gana
8	775	815	40	Cerrada	0.995	0.972	2.3%	Gana
9	841	916	75	Abierta	1.011	1.116	-9.4%	Pierde
10	917	943	26	Cerrada	1.647	1.51	9%	Gana
11	944	1044	100	Abierta	1.163	1.208	-3.7%	Pierde
12	1045	1145	100	Abierta	1.162	1.103	5.3%	Gana
13	1157	1172	15	Cerrada	1.01	0.982	2.8%	Gana
14	1176	1212	36	Cerrada	0.902	0.863	4.5%	Gana
15	1225	1313	88	Abierta	1.109	1.404	-21.0%	Pierde
16	1314	1414	100	Abierta	1.406	1.319	6.5%	Gana
17	1415	1515	100	Abierta	1.23	1.217	1.0%	Gana
18	1516	1616	100	Abierta	0.671	0.686	2.1%	Pierde
19	1617	1717	100	Abierta	1.4	1.44	-2.7%	Pierde

Cuadro 7: Resultados en pruebas con series original

5. Conclusiones

Nos preguntamos si es posible obtener excesos de retornos empleando arbitraje estadístico con tres activos y desarrollar un método que nos permita medir las posibilidades de aprovechar eventuales discrepancias.

Como respuesta a nuestra pregunta elaboramos una hipótesis en la que afirmamos que el método va a encontrar excesos de retornos en forma sistemática si y sólo si las ternas presentan reversión a un equilibrio.

Diseñamos un modelo empleando redes neuronales, que descubre la relación de equilibrio y mide la posibilidad de obtener excesos de retornos con la terna de activos presentada. El objetivo es beneficiarnos de los desequilibrios que puedan surgir y obtener excesos de retornos cuando el precio de alguno de ellos presente discrepancias con la relación de equilibrio.

Sometimos al método que desarrollamos a mil pruebas con datos simulados y verificamos que el mismo es capaz de encontrar una relación en los precios, detecta las discrepancias y cuantifica la posibilidad de éxito de la estrategia.

Finalmente, los resultados que obtuvimos representan evidencias para validar el método. En las simulaciones que presentan reversión las ganancias son sistemáticas. En el 91.6 % de las pruebas los precios convergen y dan una ganancia media igual al 7.7 %. Por otro lado en el grupo sin reversión los resultados son más volátiles y las ganancias presentan mayor aleatoriedad presentando una media total similar a los gastos de transacción. El método identifica al Grupo A como una terna para ser explotable mediante estrategias de arbitraje estadístico.

Cuando trabajamos con precios de bancos que cotizan en el mercado los resultados no son contundentes. Esto puede deberse a que estos bancos presenten algún nivel de reversión no tan fuerte. Sin embargo, si un inversor decide realizar esta estrategia con estos bancos, obtendría en promedio excesos de retornos equivalentes al 4.4 % de lo invertido. También pusimos a prueba el método con la serie original para mostrar como hubieran sido los beneficios si se hubiese invertido desde el origen de la serie. Si bien los excesos de retornos promedios son positivos (1.5 %), no parecen justificar el empleo de recursos para su seguimiento.

Los test conocidos y empleados para realizar arbitraje estadístico, son una condición

suficiente para desarrollar la estrategia, pero no los consideramos que sean necesarias. Creemos que puede usarse arbitraje estadístico en condiciones en las que los test no considerarían. Mucha literatura existe para aplicar arbitraje estadístico en pares de activos, pero muy poca se desarrolla para grupos de más activos.

Todo el desarrollo del método fue implementado con el software estadístico R cuya licencia es de libre distribución.

Futuras líneas de investigación: En el camino que recorrimos al hacer este trabajo se nos abrieron nuevas preguntas. Creemos que la respuesta a las mismas pueden ser estudiadas en futuras investigaciones.

Una línea de investigación futura inmediata podría ser diseñar un método para emplear arbitraje estadístico en grupos de activos mayores a tres. También resultaría de interés poner a prueba el método con precios intradiarios.

Como hemos visto en los resultados con datos reales, muchas veces la estrategia no se cerró pero pese a ello los excesos de retornos fueron positivos en la mayoría de estos casos. Resulta de interés entonces, definir una estrategia de salida estableciendo límites de ganancias sin necesidad de esperar a la convergencia para culminar la estrategia. Hemos observado además, como afecta la reversión al buen funcionamiento del método. Surge de interés también realizar un trabajo para sensibilizar el factor de reversión y establecer la relación que tiene con las ganancias del método.

La red neuronal que empleamos está diseñada íntegramente con funciones lineales. Es sabido que además pueden emplearse otro tipo de funciones de activación, las que pueden encontrar relaciones no lineales y por lo tanto más complejas. El empleo de este tipo de funciones necesita cierta transformación en los datos, que no fue intención en este trabajo y puede quedar pendiente su análisis y comparación para futuras investigaciones.

6. Anexo

6.1. Estructura gráfica de la red

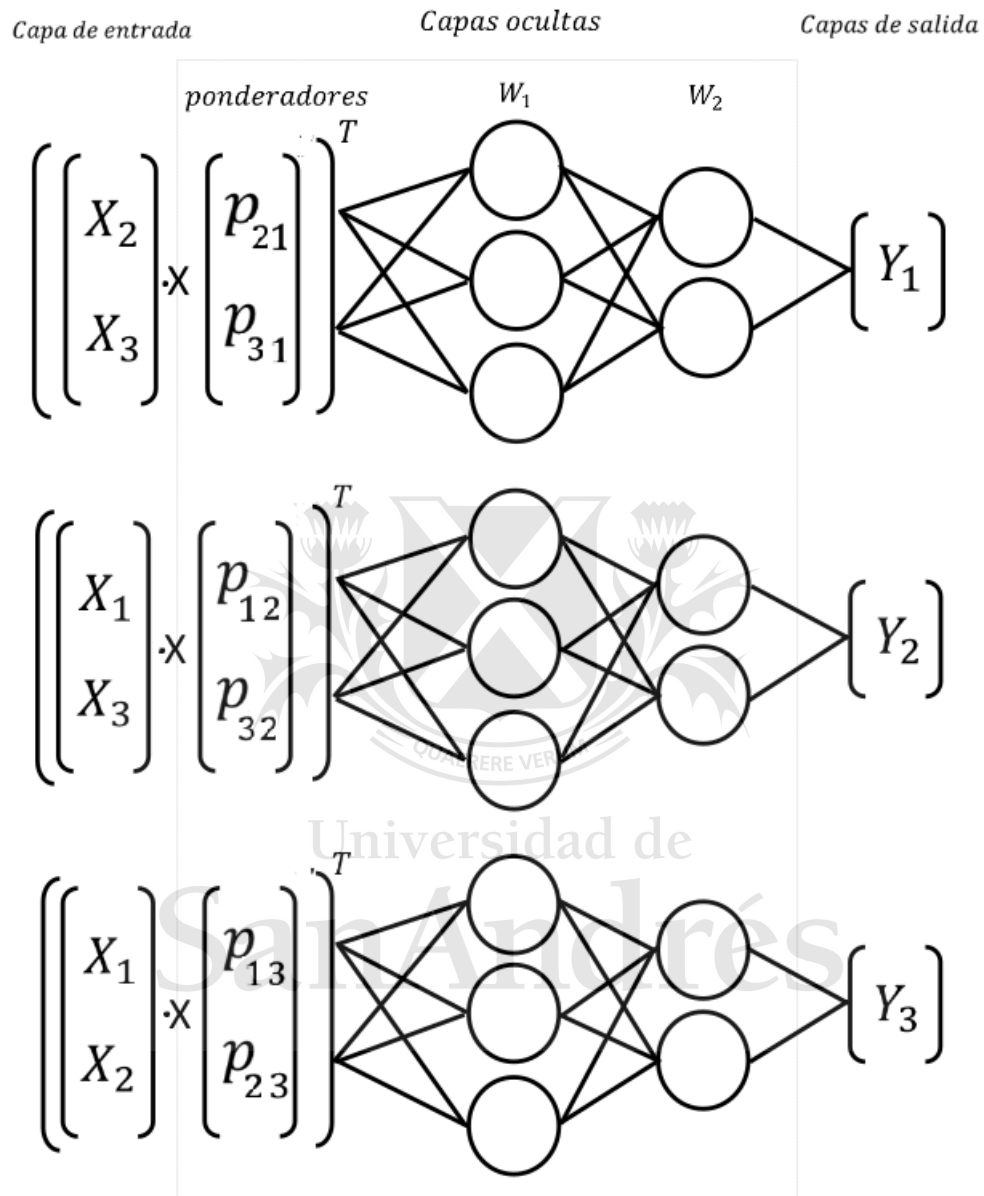


Figura 7: Estructura de la red neuronal

6.2. Estructura matricial de la red

$$g[\{f[(X * p) * W_1] * W_2\}] * W_3 = \hat{Y} \quad (1)$$

donde f y g son funciones de activación e \hat{Y} representa al activo sintético.

$$X = \begin{pmatrix} x_{2,1} & x_{3,1} & x_{1,1} & x_{3,1} & x_{1,1} & x_{2,1} \\ x_{2,2} & x_{3,2} & x_{1,2} & x_{3,2} & x_{1,2} & x_{2,2} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ x_{2,t} & x_{3,t} & x_{1,t} & x_{3,t} & x_{1,t} & x_{2,t} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ x_{2,n} & x_{3,n} & x_{1,n} & x_{3,n} & x_{1,n} & x_{2,n} \end{pmatrix} \quad (2)$$

donde cada elemento $x_{i,t}$ representa al precio del activo i en el momento t .

$$p = \begin{pmatrix} p_{2,1} & p_{3,1} & p_{1,2} & p_{3,2} & p_{1,3} & p_{2,3} \end{pmatrix} \quad (3)$$

donde cada elemento $p_{i,j}$ es el ponderador del activo i para replicar la cartera del activo j .

La red optimiza sus pesos mediante el método "patrón a patrón", es decir que ajusta los pesos observando de a un patrón a la vez. Por ejemplo para el día $t = 1$ multiplica elemento por elemento $X_{.,t=1} * p$, obteniendo una matriz de $1X6$. Esta matriz es multiplicada por W_1 . Los valores pertenecientes al mismo precio a estimar se anulan en esta multiplicación y así sucesivamente en el resto de las multiplicaciones.

$$W_1 = \begin{pmatrix} w_{1,1,1} & w_{1,2,1} & w_{1,3,1} & 0 & 0 & 0 & 0 & 0 & 0 \\ w_{2,1,1} & w_{2,2,1} & w_{2,3,1} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & w_{3,4,1} & w_{3,5,1} & w_{3,6,1} & 0 & 0 & 0 \\ 0 & 0 & 0 & w_{4,4,1} & w_{4,5,1} & w_{4,6,1} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & w_{5,7,1} & w_{5,8,1} & w_{5,9,1} \\ 0 & 0 & 0 & 0 & 0 & 0 & w_{6,7,1} & w_{6,8,1} & w_{6,9,1} \end{pmatrix} \quad (4)$$

$$W_2 = \begin{pmatrix} w_{1,1,2} & w_{1,2,2} & 0 & 0 & 0 & 0 \\ w_{2,1,2} & w_{2,2,2} & 0 & 0 & 0 & 0 \\ w_{3,1,2} & w_{3,2,2} & 0 & 0 & 0 & 0 \\ 0 & 0 & w_{4,3,2} & w_{4,4,2} & 0 & 0 \\ 0 & 0 & w_{5,3,2} & w_{5,4,2} & 0 & 0 \\ 0 & 0 & w_{6,3,2} & w_{6,4,2} & 0 & 0 \\ 0 & 0 & 0 & 0 & w_{7,5,2} & w_{7,6,2} \\ 0 & 0 & 0 & 0 & w_{8,5,2} & w_{8,6,2} \\ 0 & 0 & 0 & 0 & w_{9,5,2} & w_{9,6,2} \end{pmatrix} \quad (5)$$

$$W_3 = \begin{pmatrix} w_{1,1,3} & 0 & 0 \\ w_{2,1,3} & 0 & 0 \\ 0 & w_{3,2,3} & 0 \\ 0 & w_{4,2,3} & 0 \\ 0 & 0 & w_{5,3,3} \\ 0 & 0 & w_{6,3,3} \end{pmatrix} \quad (6)$$

donde W_k representa a los pesos de la capa k y $w_{m,n,k}$ es el elemento m de la neurona n de la capa k .

$$\hat{Y} = \begin{pmatrix} \hat{y}_{1,1} & \hat{y}_{1,2} & \hat{y}_{1,3} \\ \hat{y}_{1,2} & \hat{y}_{2,2} & \hat{y}_{3,2} \\ \vdots & \vdots & \vdots \\ \hat{y}_{1,t} & \hat{y}_{2,t} & \hat{y}_{3,t} \\ \vdots & \vdots & \vdots \\ \hat{y}_{1,n} & \hat{y}_{2,n} & \hat{y}_{3,n} \end{pmatrix} \quad (7)$$

siendo $\hat{y}_{i,t}$ el valor del sintético del activo i en el momento t .

6.3. Umbrales de equilibrio

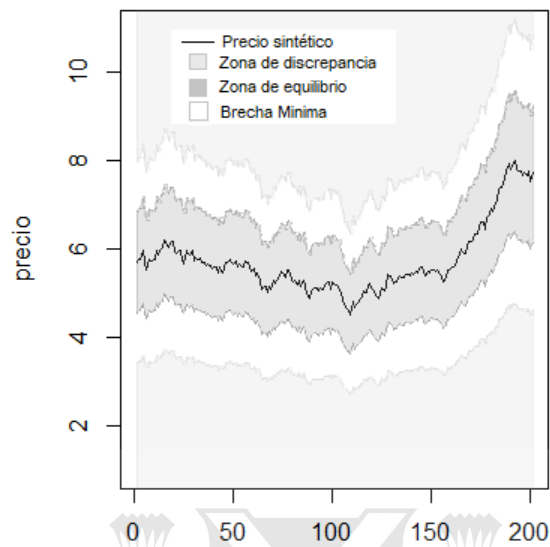


Figura 8: Umbrales y zonas de equilibrio

6.4. Ternas de precios simuladas con reversión

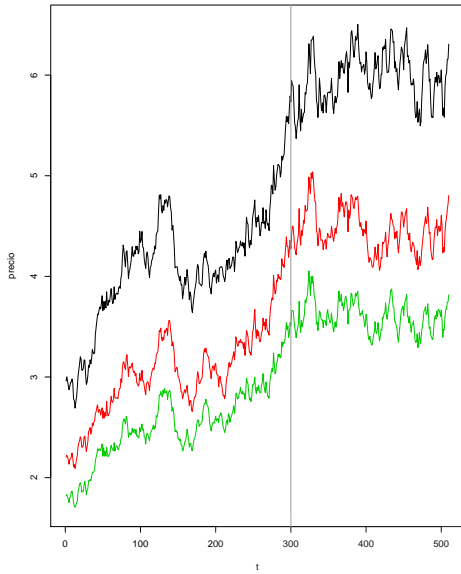


Figura 9: Simulación con reversión 1

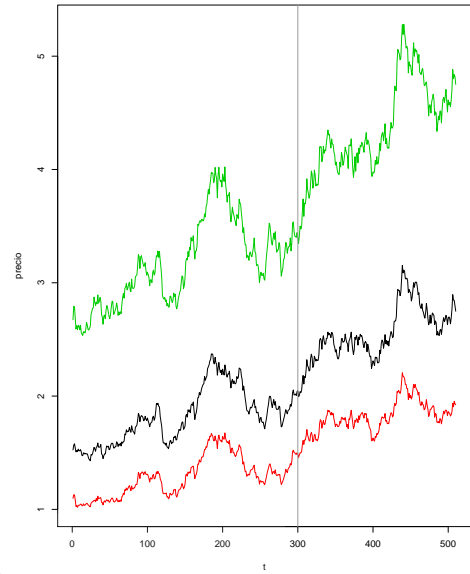


Figura 10: Simulación con reversión 2

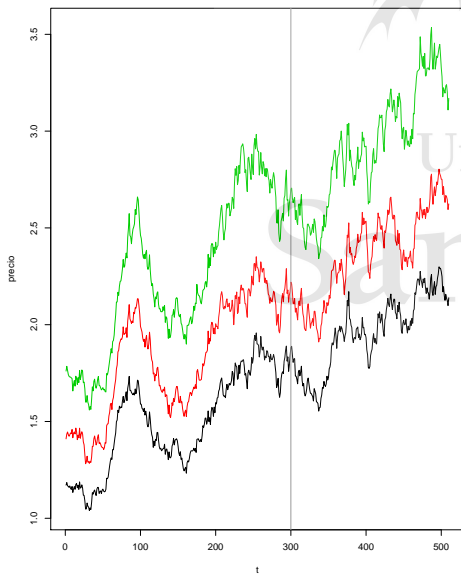


Figura 11: Simulación con reversión 3

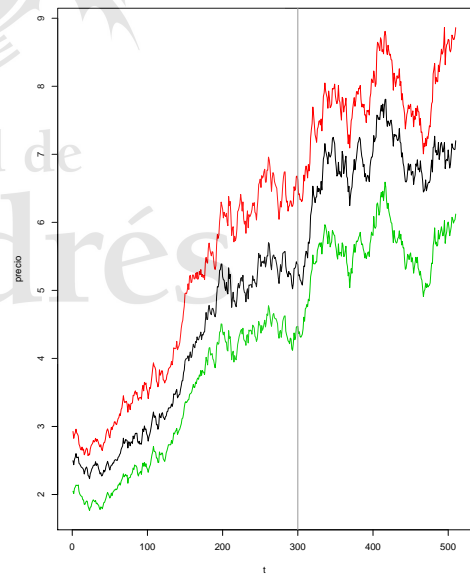


Figura 12: Simulación con reversión 4

6.5. Ternas de precios simuladas sin reversión

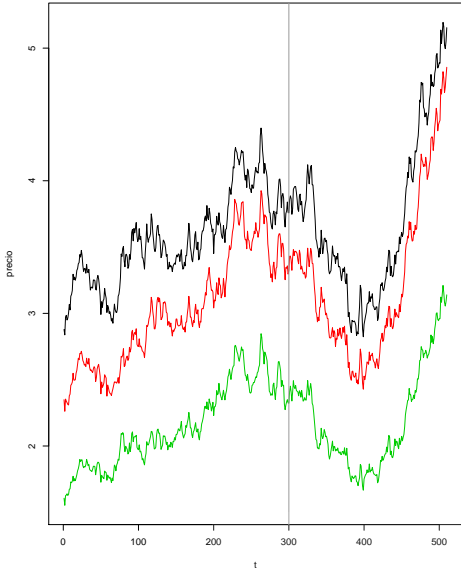


Figura 13: Simulación sin reversión 1

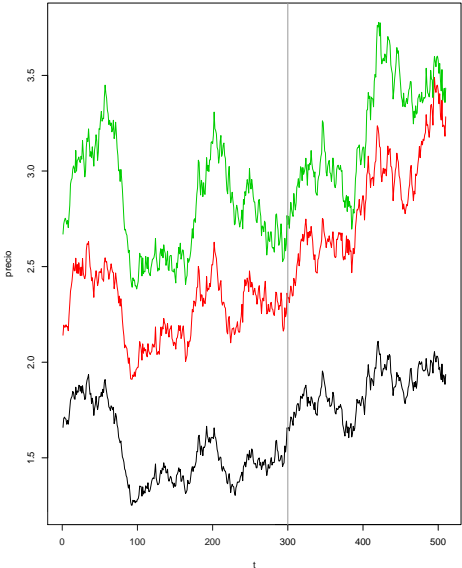


Figura 14: Simulación sin reversión 2

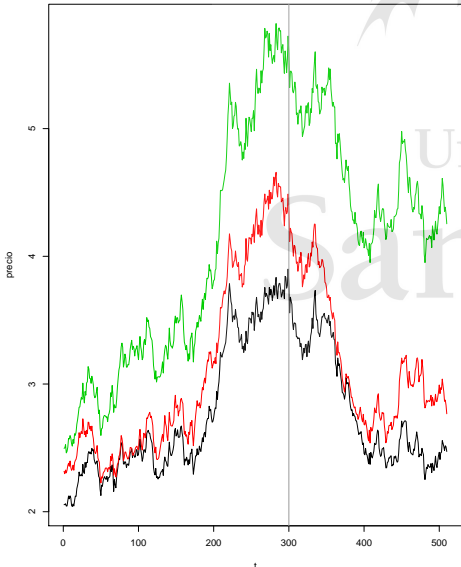


Figura 15: Simulación sin reversión 3

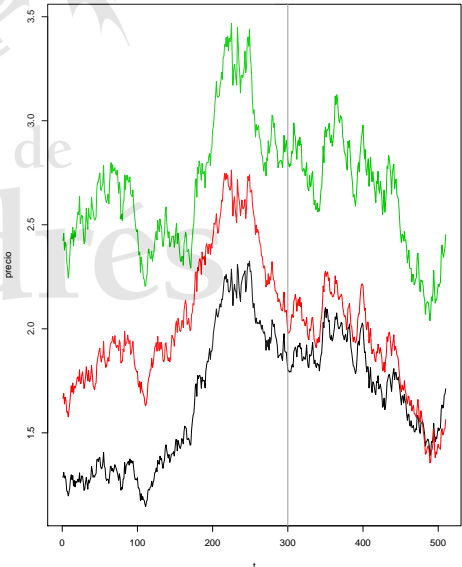


Figura 16: Simulación sin reversión 4

6.6. Ternas de precios simuladas con datos reales

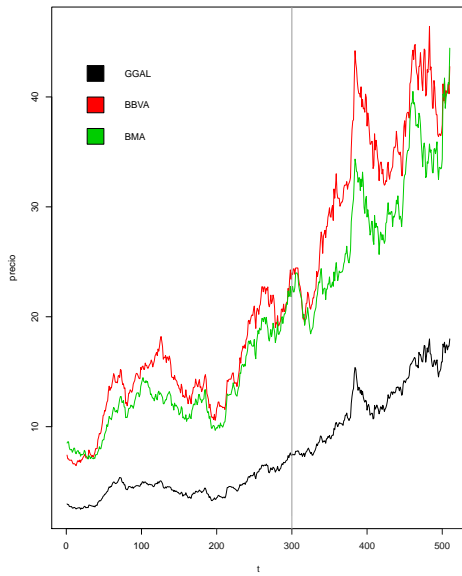


Figura 17: Simulación de bancos 1

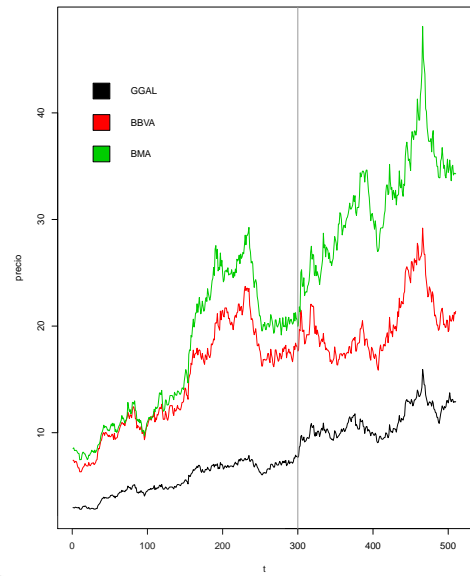


Figura 18: Simulación de bancos 2

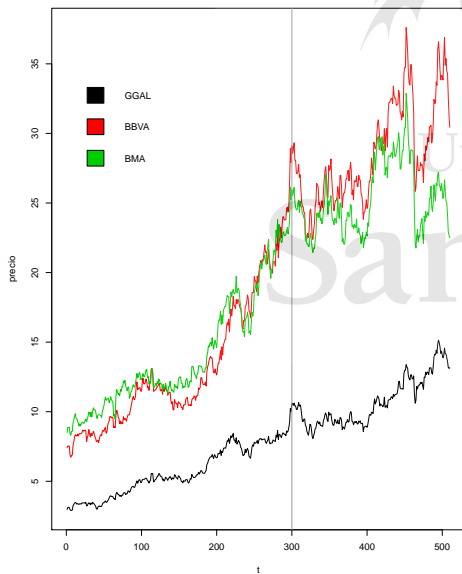


Figura 19: Simulación de bancos 3

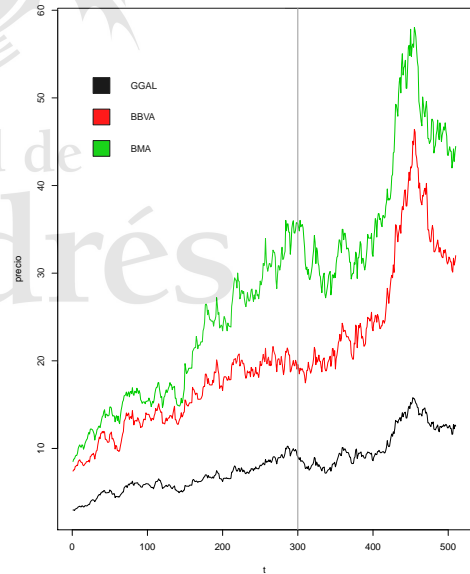


Figura 20: Simulación de bancos 4

6.7. Tabla de resultados con datos simulados con reversión

Simulación	Compra	Venta	Inicio	Final	t	Ex.Ret.	Ex.Ret. Anual	Estado	Ganancia	p-value
1	1.194	1.088	324	432	108	0.097	0.240	Cerrada	Gana	0.45
2	1.739	1.692	313	500	187	0.028	0.037	Abierta	Gana	0.76
3	1.254	1.190	336	384	48	0.053	0.311	Cerrada	Gana	0.82
4	1.602	1.452	301	500	199	0.102	0.130	Abierta	Gana	0.05
5	1.621	1.406	301	496	195	0.153	0.200	Cerrada	Gana	0.54
6	1.017	0.978	304	363	59	0.039	0.179	Cerrada	Gana	0.13
7	1.208	1.094	301	331	30	0.103	1.273	Cerrada	Gana	0.44
8	1.338	1.345	305	500	195	-0.00	-0.00	Abierta	Pierde	0.30
9	1.429	1.372	329	392	63	0.041	0.174	Cerrada	Gana	0.37
10	0.972	0.910	301	349	48	0.067	0.405	Cerrada	Gana	0.47
11	0.964	0.891	301	312	11	0.081	4.978	Cerrada	Gana	0.68
12	1.046	0.996	307	354	47	0.049	0.296	Cerrada	Gana	0.07
13	0.986	0.924	305	331	26	0.067	0.877	Cerrada	Gana	0.17
14	1.457	1.358	303	407	104	0.073	0.184	Cerrada	Gana	0.37
15	1.147	0.997	302	342	40	0.150	1.404	Cerrada	Gana	0.52
16	1.080	1.027	301	336	35	0.051	0.429	Cerrada	Gana	0.01
17	0.944	0.877	323	343	20	0.076	1.524	Cerrada	Gana	0.01
18	1.043	1.133	302	500	198	-0.07	-0.09	Abierta	Pierde	0.18
19	1.161	1.068	301	338	37	0.086	0.753	Cerrada	Gana	0.53
20	1.671	1.515	311	349	101	0.103	0.274	Cerrada	Gana	0.30
21	1.063	1.011	312	457	40	0.050	0.361	Cerrada	Gana	0.01
22	0.890	0.845	301	402	66	0.053	0.219	Cerrada	Gana	0.11
23	1.051	0.957	320	360	36	0.097	0.911	Cerrada	Gana	0.01
24	1.167	1.093	307	373	49	0.067	0.398	Cerrada	Gana	0.08
25	1.066	0.975	301	337	43	0.092	0.675	Cerrada	Gana	0.01
26	1.173	1.082	301	350	19	0.084	1.904	Cerrada	Gana	0.56
27	1.038	0.953	301	344	6	0.089	33.94	Cerrada	Gana	0.31
28	1.189	1.139	301	320	145	0.044	0.078	Cerrada	Gana	0.06
29	1.360	1.203	316	322	128	0.130	0.271	Cerrada	Gana	0.82
30	1.842	1.614	317	462	195	0.141	0.185	Cerrada	Gana	0.38
31	1.372	1.279	301	429	32	0.072	0.726	Cerrada	Gana	0.91
32	1.208	1.129	305	500	48	0.069	0.416	Abierta	Gana	0.01
33	1.092	1.015	307	339	26	0.076	1.032	Cerrada	Gana	0.63
34	1.059	0.966	301	349	33	0.096	1.004	Cerrada	Gana	0.01
35	1.166	1.047	308	334	59	0.113	0.579	Cerrada	Gana	0.64
36	1.132	1.065	301	334	43	0.063	0.426	Cerrada	Gana	0.39
37	1.596	1.492	301	360	167	0.069	0.106	Cerrada	Gana	0.01
38	1.318	1.230	310	353	50	0.071	0.415	Cerrada	Gana	0.80
39	0.891	0.780	301	468	68	0.142	0.630	Cerrada	Gana	0.10
40	0.990	0.934	312	362	28	0.060	0.688	Cerrada	Gana	0.65

Cuadro 8: Simulaciones con reversión

6.8. Tabla de resultados con datos simulados sin reversión

Simulación	Compra	Venta	Inicio	Final	t	Ex.Ret.	Ex.Ret. Anual	Estado	Ganancia	p-value
1	1.539	1.159	301	500	199	0.328	0.428	Abierta	Gana	0.74
2	1.242	1.314	301	500	199	-0.05	-0.06	Abierta	Pierde	0.99
3	1.235	1.467	319	500	181	-0.15	-0.21	Abierta	Pierde	0.37
4	1.055	0.992	301	500	199	0.064	0.081	Abierta	Gana	0.84
5	1.169	1.335	328	500	172	-0.12	-0.17	Abierta	Pierde	0.38
6	0.952	0.888	303	334	31	0.071	0.743	Cerrada	Gana	0.75
7	0.974	0.892	301	312	11	0.091	6.297	Cerrada	Gana	0.60
8	1.519	1.200	301	500	199	0.265	0.344	Abierta	Gana	0.85
9	1.450	1.698	318	500	182	-0.14	-0.19	Abierta	Pierde	0.97
10	1.949	1.405	301	500	199	0.387	0.508	Abierta	Gana	0.87
11	1.276	1.197	308	348	40	0.065	0.488	Cerrada	Gana	0.90
12	1.861	1.582	301	500	199	0.176	0.226	Abierta	Gana	0.09
13	1.585	1.676	318	500	182	-0.05	-0.07	Abierta	Pierde	0.53
14	1.349	1.359	301	500	199	-0.00	-0.00	Abierta	Pierde	0.19
15	1.813	1.673	301	500	199	0.083	0.105	Abierta	Gana	0.72
16	1.553	1.841	315	500	185	-0.15	-0.20	Abierta	Pierde	0.67
17	1.176	1.328	301	500	199	-0.11	-0.14	Abierta	Pierde	0.93
18	1.175	1.062	302	313	11	0.105	8.772	Cerrada	Gana	0.05
19	1.357	0.981	301	500	199	0.383	0.503	Abierta	Gana	0.65
20	1.085	1.404	309	500	135	-0.22	-0.37	Abierta	Pierde	0.96
21	1.609	1.610	301	500	199	-0.00	-0.00	Abierta	Pierde	0.85
22	0.827	1.253	365	500	180	-0.33	-0.43	Abierta	Pierde	0.77
23	1.203	1.305	301	500	192	-0.07	-0.10	Abierta	Pierde	0.52
24	1.039	0.996	320	500	18	0.043	0.799	Abierta	Gana	0.93
25	1.036	0.961	308	500	25	0.077	1.112	Abierta	Gana	0.93
26	1.615	1.722	302	320	197	-0.06	-0.07	Cerrada	Pierde	0.01
27	1.530	1.014	302	327	195	0.508	0.693	Cerrada	Gana	0.84
28	1.453	1.603	303	500	199	-0.09	-0.11	Abierta	Pierde	0.59
29	1.612	1.504	305	500	141	0.071	0.130	Abierta	Gana	0.99
30	1.056	1.365	301	500	198	-0.22	-0.27	Abierta	Pierde	0.50
31	1.186	0.914	301	442	199	0.296	0.385	Cerrada	Gana	0.60
32	1.084	1.583	302	500	179	-0.31	-0.41	Abierta	Pierde	0.72
33	1.152	1.451	301	500	199	-0.20	-0.25	Abierta	Pierde	0.60
34	1.335	1.465	321	500	199	-0.08	-0.11	Abierta	Pierde	0.56
35	1.477	1.316	301	500	184	0.122	0.170	Abierta	Gana	0.95
36	1.216	1.032	301	500	59	0.178	1.002	Abierta	Gana	0.33
37	1.357	2.039	303	487	199	-0.33	-0.39	Cerrada	Pierde	0.63
38	1.395	1.485	301	360	199	-0.06	-0.07	Cerrada	Pierde	0.27
39	1.308	1.350	301	500	189	-0.03	-0.04	Abierta	Pierde	0.49
40	1.082	1.308	301	500	199	-0.17	-0.21	Abierta	Pierde	0.02

Cuadro 9: Simulaciones sin reversión

6.9. Tabla de resultados con datos simulados de bancos

Simulación	Compra	Venta	Inio	Final	t	Ex.Ret.	Ex.Ret. Anual	Estado	Ganancia	p-value
1	3.121	3.140	311	500	189	-0.00	-0.00	Abierta	Pierde	0.94
2	0.914	0.842	305	500	195	0.085	0.110	Abierta	Gana	0.90
3	1.454	1.285	301	370	69	0.131	0.566	Cerrada	Gana	0.99
4	0.996	0.918	301	341	40	0.084	0.661	Cerrada	Gana	0.60
5	3.606	2.630	304	500	196	0.371	0.495	Abierta	Gana	0.99
6	1.564	1.773	303	500	197	-0.11	-0.14	Abierta	Pierde	0.55
7	1.445	1.700	319	500	181	-0.15	-0.20	Abierta	Pierde	0.88
8	2.326	2.219	303	500	197	0.048	0.061	Abierta	Gana	0.78
9	1.830	1.630	303	500	197	0.122	0.158	Abierta	Gana	0.30
10	0.834	1.131	301	500	199	-0.26	-0.31	Abierta	Pierde	0.09
11	1.610	1.323	305	500	195	0.217	0.286	Abierta	Gana	0.75
12	2.549	2.136	301	500	199	0.193	0.248	Abierta	Gana	0.94
13	1.138	1.005	301	453	152	0.132	0.226	Cerrada	Gana	0.01
14	1.559	1.684	304	500	196	-0.07	-0.09	Abierta	Pierde	0.92
15	1.233	1.633	301	500	199	-0.24	-0.29	Abierta	Pierde	0.01
16	1.264	1.177	484	499	15	0.074	2.292	Cerrada	Gana	0.06
17	1.094	1.265	301	500	199	-0.13	-0.16	Abierta	Pierde	0.37
18	3.191	2.461	308	500	192	0.296	0.402	Abierta	Gana	0.61
19	1.073	0.978	309	339	30	0.096	1.149	Cerrada	Gana	0.89
20	1.075	1.010	305	322	27	0.064	0.785	Cerrada	Gana	0.74
21	1.505	1.474	318	372	158	0.020	0.032	Cerrada	Gana	0.22
22	0.708	0.948	301	328	178	-0.25	-0.33	Cerrada	Pierde	0.62
23	1.293	1.097	342	500	173	0.178	0.268	Abierta	Gana	0.44
24	0.835	0.779	322	500	34	0.071	0.665	Abierta	Gana	0.19
25	1.427	1.318	301	474	107	0.082	0.204	Cerrada	Gana	0.08
26	1.547	1.223	301	335	49	0.264	2.316	Cerrada	Gana	0.91
27	1.288	1.517	310	417	199	-0.15	-0.18	Cerrada	Pierde	0.95
28	1.747	1.615	301	350	199	0.081	0.103	Cerrada	Gana	0.30
29	2.100	1.790	301	500	62	0.173	0.905	Abierta	Gana	0.27
30	1.047	0.947	301	500	189	0.105	0.142	Abierta	Gana	0.86
31	2.125	1.879	301	363	121	0.131	0.290	Cerrada	Gana	0.30
32	1.281	1.854	301	490	186	-0.30	-0.39	Cerrada	Pierde	0.50
33	1.651	1.792	301	422	195	-0.07	-0.09	Cerrada	Pierde	0.90
34	1.671	1.555	314	500	199	0.075	0.095	Abierta	Gana	0.73
35	1.431	1.515	305	500	199	-0.05	-0.06	Abierta	Pierde	0.38
36	1.890	1.684	301	500	86	0.121	0.397	Abierta	Gana	0.22
37	1.082	1.048	301	500	15	0.032	0.708	Abierta	Gana	0.48
38	1.677	1.593	323	409	120	0.052	0.113	Cerrada	Gana	0.42
39	1.613	1.481	301	316	166	0.088	0.136	Cerrada	Gana	0.95
40	0.990	0.876	313	433	20	0.130	3.618	Cerrada	Gana	0.04

Cuadro 10: Simulaciones con precios reales

7. Bibliografía

Referencias

- [1] Lo, Andrew; MacKinlay, Craig (1988). Stock Market Prices Do Not Follow Random Walks: Evidence from a Simple Specification Test. *The Review of Financial Studies*, 1 (1), pp. 41-66.
- [2] Guidolin, Massimo; Hyde, Stuart; McMillan, David; Ono, Sadayuki (2009). Non-linear predictability in stock and bond returns: When and where is it exploitable?. *International Journal of Forecasting*, 25, pp. 373-399.
- [3] Barberis, N.; Shleifer, A.; Wurgler, J. (2005). Comovement, *Journal of Financial Economics*, 75, pp. 283-317.
- [4] Triantafyllopoulos, Kostas ; Montana, Giovanni (2009). Dynamic modeling of mean-reverting spreads for statistical arbitrage.
- [5] Bustamante, S. (22 de noviembre de 2017). Vuelven a habilitar las ventas en corto para darle liquidez al mercado. *El cronista*. <https://www.cronista.com/finanzasmercados/Vuelven-a-habilitar-las-ventas-en-corto-para-darle-liquidez-al-mercado-20171122-0034.html>
- [6] Bezán, A. (17 de abril de 2018). Arrancan en una semana las ventas en corto. *Clarín*. <https://www.clarin.com/economia/arrancan-semana-ventas-corto.html>
- [7] Fama, E. F. (1970), Efficient Capital Markets: A Review of Theory and Empirical Work. *Journal of Finance*, 25 (2), Papers and Proceedings of the Twenty-Eighth Annual Meeting of the American Finance Association, New York, N.Y. pp. 383-417.
- [8] Haugen, R. A. (1999), *The New Finance: The Case Against Efficient Markets*, Prentice Hall.
- [9] Los, C. A. (2000), Nonparametric efficiency testing of Asian markets using weekly data. *Advances in Econometrics*, 14, pp. 329-363.
- [10] Box, G. E. P., Jenkins, G. M., and Reinsel, C. (1970). *Forecasting and control*. *Times Series Analysis*.

- [11] Cybenko G (1989). Approximations by Superpositions of a Sigmoidal Function. *Mathematics of Control, Signals, and Systems*. 2: 303-314.
- [12] Avellaneda, M.; Lee, J. (2010). Statistical arbitrage in the US equities market, pp. 1-22.
- [13] Burgess, N.; Refenes, A. N. (1996), Modelling nonlinear cointegration in international equity index futures, *Neural Networks in Financial Engineering*, World Scientific, pp. 50-63.
- [14] Shadbolt, J.; Taylor, J. G. (2002). *Cointegration, Neural Networks and the Financial Markets: predicting, combining and portfolio optimisation*, Springer, pp. 181-191.
- [15] Towers, N. (2002), Joint optimisation in statistical arbitrage trading, *Neural Networks and the Financial Markets: predicting, combining and portfolio optimisation*, Springer, pp. 193-201.
- [16] Garrett, I. and Taylor, N. (2001), Intraday and interday basis dynamics: Evidence from the FTSE 100 index futures market, *Studies in Nonlinear Dynamics and Econometrics*, 5 (2), pp. 133-152.
- [17] Steurer, E.; Hann, T. H. (1996), Exchange rate forecasting comparison: neural networks, machine learning and linear models, *Neural Networks in Financial Engineering*, World Scientific, pp. 113-121.
- [18] W. Enders (1995). *Applied econometric analysis*, Jhon Wiley and Sons.
- [19] Gatev, E.; Goetzmann, W.N.; Rowenhorst, K.G (2006). Pairs trading: Performance of a relative value arbitrage rule. *Review of Financial Studies*, v. 19, p. 797-827.
- [20] Andrade, C.S.; DI Pietro, V.; Seasholes, M.S (2005). Understanding the profitability of pairs trading. Technical report, UC Berkeley Haas School.
- [21] Engelberg, J.; Gao, P.; Jagannathan, R. (2009). An anatomy of pairs trading: the role of idiosyncratic news, common information and liquidity. In *Proceedings of the Third Singapore International Conference on Finance*.
- [22] Do, B.; Faff, R.W (2010). Does simple pairs trading still work? *Financial Analysts Journal*, 66, pp. 83-95.

- [23] Bowen, D.; Hutchinson, M.; Sullivan, N. (2010). High frequency equity pairs trading: transaction costs, speed of execution and patterns in returns. *Journal of Trading*, 5, pp. 31-38,
- [24] Broussard, J. P.; Vaihekoski M. (2012). Profitability of pairs trading strategy in an illiquid market with multiple share classes. *Journal of International Financial Markets, Institutions and Money*, 22, pp. 1188-1201.
- [25] Lucas, A. (1997). Strategic and tactical asset allocation and the effect of long-run equilibrium relations. *Serie Research Memoranda 0042*, VU University Amsterdam.
- [26] Alexander, C. (1999). Optimal hedging using cointegration. *Philosophical Transactions of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences*, 357, pp. 2039-2058.
- [27] Alexander, C.; Dimitru, A. (2002): The cointegration alpha: Enhanced index tracking and longshort equity market neutral strategies.
- [28] Bessler, D. A.; Yang, J. (2003). The structure of interdependence in international stock markets. *Journal of International Money and Finance*, v. 22, pp. 261-287.
- [29] Yang, J.; Kolari, J.W.; Sutanto, P.W. (2004). On the stability of long-run relationships between emerging and US stock markets. *Journal of Multinational Financial Management*, 14, pp. 233-248.
- [30] Engle, R.F.; Granger, C.W. (1987). Co-integration and error correction: representation, estimation, and testing. *Econometrica*, pp. 251-276.
- [31] Johansen, S. (1988) Statistical analysis of cointegration vectors, *Journal of Economic Dynamics and Control*, 12, pp. 231-54.
- [32] Elliott, R., van der Hoek, J. and Malcolm, W. (2005) Pairs Trading, *Quantitative Finance*, 5 (3), pp. 271-276.
- [33] Caldas, Bruno and Caldeira Joao (2006). Pairs Trading: Different Weights and Markets.