

Estrategias Dinámicas de Pairs Trading:

Aplicación de Filtros de Kalman en Cointegración Adaptativa

José de Jesús Morales Domínguez

Proyecto 4 Trading

Resumen Ejecutivo

Este reporte presenta una investigación exhaustiva sobre la aplicación de Filtros de Kalman en estrategias de Pairs Trading basadas en cointegración dinámica. A diferencia de los enfoques tradicionales que asumen hedge ratios constantes, nuestra metodología implementa parámetros adaptativos que se ajustan en tiempo real a las condiciones cambiantes del mercado.

El par Johnson & Johnson (JNJ) - Pfizer (PFE) demostró una relación de cointegración robusta fundamentada en características económicas compartidas del sector farmacéutico, permitiendo la implementación exitosa de una estrategia market-neutral con métricas de riesgo-retorno excepcionales.

Los resultados en el conjunto de entrenamiento muestran un retorno anualizado de 37.92% con un Sharpe Ratio de 0.97 y un win rate de 88.89%. El máximo drawdown observado fue de apenas 1.03%, resultando en un Calmar Ratio de 36.66. El Profit Factor alcanzó 7846.51, indicando una relación extremadamente favorable entre ganancias y pérdidas.

En el conjunto de test, que representa datos completamente no vistos, la estrategia demostró notable robustez con un retorno anualizado de 46.07%, manteniendo el Sharpe Ratio en 0.97 y mejorando el Sortino a 9.23. El máximo drawdown se redujo a 0.49%, elevando el Calmar Ratio a 94.34. El win rate fue de 66.67% con un Profit Factor de 234.08.

Introducción

El trading de activos individuales presenta desafíos fundamentales que han sido ampliamente documentados en la literatura financiera. Los precios de acciones siguen un camino aleatorio según la Random Walk Hypothesis propuesta por Malkiel en 1973, haciendo que los cambios de precio sean impredecibles y difíciles de modelar sistemáticamente. Adicionalmente, los traders enfrentan exposición al riesgo de mercado (beta), donde movimientos sistémicos pueden invalidar análisis fundamentales correctos. La no-estacionariedad de los precios viola supuestos básicos de muchos modelos estadísticos y de machine learning, mientras que el

fenómeno de data drift hace que los patrones históricos rápidamente se vuelvan obsoletos.

Las estrategias de Pairs Trading ofrecen una solución elegante a estos problemas al enfocarse en relaciones relativas entre activos en lugar de movimientos absolutos. Al crear posiciones market-neutral que minimizan la exposición al riesgo sistemático, estas estrategias explotan arbitraje estadístico basado en relaciones de equilibrio económico. La clave está en operar sobre spreads estacionarios que exhiben reversión a la media, transformando el problema del trading direccional en uno de identificación de desviaciones temporales de equilibrio.

El Pairs Trading fue popularizado en los años 1980s por analistas cuantitativos de Morgan Stanley, particularmente Gerry Bamberger y Nunzio Tartaglia. Inicialmente desarrollado como una estrategia de arbitraje estadístico puro, el enfoque ha evolucionado significativamente a través de las décadas. Durante los años 1990s se introdujeron métodos formales de cointegración basados en los trabajos de Engle y Granger, así como el test multivariado de Johansen. La década de 2000 vio la incorporación de modelos dinámicos y técnicas de machine learning, mientras que en los años 2010 hasta el presente, el trading de alta frecuencia, deep learning y filtros adaptativos han dominado el desarrollo de estas estrategias.

Los métodos clásicos de Pairs Trading basados en Engle-Granger o Johansen calculan el hedge ratio β mediante regresión OLS sobre datos históricos, asumiendo que esta relación permanece constante a lo largo del tiempo. El modelo estático tradicional se expresa como $Y_t = \beta_0 + \beta_1 \cdot X_t + v_t$, donde β_0 y β_1 son parámetros fijos estimados una sola vez.

Esta asunción de estacionariedad en los parámetros presenta problemas críticos en la práctica. Los mercados experimentan cambios de régimen asociados a crisis financieras, cambios regulatorios y eventos macroeconómicos que alteran fundamentalmente las relaciones entre activos. Las relaciones económicas entre empresas evolucionan con el tiempo debido a cambios en estrategias corporativas, fusiones y adquisiciones, y desarrollo de nuevos productos. El overfitting a datos históricos (data snooping) no garantiza performance futura, y los modelos estáticos carecen de mecanismos para responder dinámicamente a nuevas condiciones de mercado.

Este reporte presenta una solución mediante Filtros de Kalman, que permiten parámetros time-varying donde $\beta_0(t)$ y $\beta_1(t)$ se actualizan dinámicamente con cada nueva observación. El enfoque constituye un algoritmo de online learning donde el modelo aprende continuamente de nuevos datos sin necesidad de re-entrenamiento

completo. El filtro mantiene estimaciones de incertidumbre en tiempo real, cuantificando la confianza en cada estimación de parámetros. Esta adaptabilidad automática permite que el sistema responda eficientemente a cambios de régimen de mercado sin intervención manual.

Marco Teórico

Estacionariedad

Una serie temporal es estacionaria si sus propiedades estadísticas no cambian con el tiempo. Formalmente, esto requiere que la media sea constante para todo t , es decir $E[X_t] = \mu$ para cualquier punto en el tiempo. La varianza también debe permanecer constante, $\text{Var}(X_t) = \sigma^2$ independientemente de t . Finalmente, la autocovarianza debe depender únicamente del lag entre observaciones, no del punto específico en el tiempo, expresado como $\text{Cov}(X_t, X_{t-k}) = \gamma_k$.

La estacionariedad es crucial en trading cuantitativo porque muchos modelos estadísticos, incluyendo regresión por mínimos cuadrados ordinarios, modelos ARIMA y numerosos algoritmos de machine learning, asumen esta propiedad. Los precios de acciones son típicamente no estacionarios, siguiendo procesos de random walk. Sin embargo, los retornos diarios y ciertos spreads de cointegración pueden exhibir estacionariedad, permitiendo la aplicación de técnicas estadísticas tradicionales.

Procesos Integrados

Un proceso X_t es integrado de orden d , denotado $I(d)$, si requiere d diferenciaciones para convertirse en estacionario. Matemáticamente, esto significa que $\Delta^d X_t$ es estacionario, mientras que $\Delta^{d-1} X_t$ no lo es. Los procesos $I(0)$ son series estacionarias como ruido blanco o retornos de activos. Los procesos $I(1)$, como random walks, requieren una diferenciación y representan la mayoría de series de precios financieros. Los procesos $I(2)$, que requieren dos diferenciaciones, son raros en finanzas pero pueden aparecer en datos macroeconómicos.

Test de Dickey-Fuller Aumentado

El test ADF (Augmented Dickey-Fuller) es la herramienta estándar para detectar raíces unitarias en series temporales. La hipótesis nula establece que la serie tiene raíz unitaria y por tanto es no estacionaria, mientras que la hipótesis alternativa indica que la serie es estacionaria. La ecuación del test se expresa como $\Delta Y_t = \alpha + \beta t + \gamma Y_{t-1} + \sum \delta_i \Delta Y_{t-i} + \varepsilon_t$, donde los rezagos adicionales controlan por autocorrelación.

La interpretación del test es directa. Si el p-value es menor a 0.05, rechazamos la hipótesis nula y concluimos que la serie es estacionaria. Si el p-value excede 0.05, no podemos rechazar la hipótesis nula y la serie se considera no estacionaria. Este test es fundamental para verificar las condiciones necesarias de cointegración.

Cointegración: Definición Formal

Dos series X_t y Y_t son cointegradas si cumplen dos condiciones esenciales. Primero, ambas series deben ser individualmente $I(1)$, es decir, no estacionarias en niveles pero estacionarias en primeras diferencias. Segundo, debe existir una combinación lineal de estas series que sea $I(0)$, es decir estacionaria. Esta combinación se expresa como $Z_t = Y_t - \beta_1 X_t - \beta_0 \sim I(0)$.

La interpretación económica es profunda. Aunque X_t y Y_t pueden desviarse temporalmente, están vinculadas por una relación de equilibrio de largo plazo. Las desviaciones del equilibrio, representadas por Z_t , son temporales y exhiben reversión a la media. Esta propiedad es la base fundamental del Pairs Trading.

Una analogía útil es la "pareja de borrachos" descrita en la literatura de econometría. Imaginemos dos personas saliendo de un bar. La persona X camina aleatoriamente siguiendo un random walk. La persona Y también camina aleatoriamente, pero intenta mantenerse cerca de X, quizás tomados de la mano. Individualmente, ambos siguen procesos $I(1)$ con trayectorias impredecibles. Sin embargo, la distancia entre ellos es $I(0)$, mostrando reversión a la media. Esta relación es cointegración en su forma más intuitiva.

Cointegración y Pairs Trading

Correlación vs Cointegración

Es crucial distinguir entre correlación y cointegración, conceptos frecuentemente confundidos en la práctica. La correlación es una medida de corto plazo que captura movimientos sincronizados entre dos series durante un período específico. Puede cambiar rápidamente con las condiciones de mercado y genera señales de trading ruidosas cuando se usa como único criterio.

La cointegración, en contraste, captura una relación estructural de largo plazo. Representa un equilibrio económico fundamental que persiste a través de diferentes condiciones de mercado. Esta estabilidad estructural proporciona una base confiable para estrategias de reversión a la media.

Para ilustrar la diferencia, consideremos dos series con correlación alta pero sin cointegración. Ambas series podrían mostrar tendencias alcistas simultáneas (Serie A: $\uparrow\uparrow\uparrow$, Serie B: $\uparrow\uparrow\uparrow$), generando correlación positiva cercana a 1. Sin embargo, si no existe un mecanismo de equilibrio, las series pueden diverger indefinidamente, haciendo imposible una estrategia de trading basada en reversión. Por el contrario, dos series cointegradas pueden mostrar correlaciones variables en el corto plazo, pero mantienen una relación de equilibrio estable que garantiza convergencia eventual, creando oportunidades de trading explotables.

Test de Engle-Granger

El test de Engle-Granger es un procedimiento de dos pasos para verificar cointegración entre dos series temporales. El primer paso consiste en estimar una regresión OLS: $Y_t = \beta_0 + \beta_1 X_t + u_t$. Los residuos estimados \hat{u}_t representan las desviaciones del equilibrio de largo plazo. El parámetro β_1 obtenido se conoce como hedge ratio y representa la relación óptima para construir una posición market-neutral.

El segundo paso aplica el test ADF a los residuos. Si los residuos son estacionarios ($p\text{-value} < 0.05$), concluimos que las series originales están cointegradas. Este resultado implica que aunque Y_t y X_t puedan desviarse temporalmente, existe una fuerza restauradora que las mantiene ligadas en el largo plazo.

Las ventajas del método incluyen su simplicidad de implementación y su efectividad con dos series. Sin embargo, presenta limitaciones importantes. Solo puede detectar una relación de cointegración y no se extiende naturalmente a múltiples series. Más relevante para nuestro trabajo, asume que β_1 permanece constante, una limitación que abordamos mediante Filtros de Kalman.

Test de Johansen

El método de Johansen es una generalización multivariada que permite analizar cointegración entre múltiples series simultáneamente. El enfoque se basa en un modelo VAR (Vector Autoregression) expresado en forma de corrección de error: $\Delta X_t = \Pi X_{t-1} + \Gamma \Delta X_{t-1} + \dots + \varepsilon_t$. La matriz Π contiene la información sobre cointegración y puede descomponerse como $\Pi = \alpha\beta'$, donde las columnas de β son los vectores de cointegración.

El test de Johansen genera dos estadísticos principales. El Trace Statistic prueba la hipótesis nula de que existen a lo más r relaciones de cointegración contra la

alternativa de más de r relaciones. El Max Eigenvalue Statistic prueba la hipótesis de exactamente r relaciones contra $r+1$ relaciones.

Las ventajas sobre Engle-Granger son significativas. El método puede detectar múltiples relaciones de cointegración simultáneamente, permite analizar portfolios de N activos, y es más robusto estadísticamente al estimar el sistema completo en un solo paso. Los eigenvectores β resultantes definen las combinaciones lineales estacionarias que pueden usarse para construir spreads de trading.

Condiciones Necesarias para Pairs Trading

Para que una estrategia de pares sea viable en la práctica, deben cumplirse varias condiciones fundamentales. Debe existir una relación económica subyacente basada en fundamentals compartidos, como pertenecer al mismo sector, compartir cadenas de suministro, o responder a drivers macroeconómicos comunes. Los precios individuales deben ser no estacionarios, verificable mediante el test ADF con p-values mayores a 0.05.

El spread construido debe ser estacionario, confirmado mediante ADF en residuos con p-value menor a 0.05. Crucialmente, la relación debe persistir out-of-sample en datos no vistos durante el entrenamiento, demostrando que no es un artefacto de overfitting. Finalmente, ambos activos deben tener liquidez suficiente para permitir ejecución eficiente sin impacto significativo en precios o costos de transacción prohibitivos.

Filtros de Kalman

Fundamentos del Filtro de Kalman

El Filtro de Kalman es un algoritmo recursivo de estimación óptima Bayesiana desarrollado originalmente por Rudolf Kalman en 1960. Originalmente diseñado para problemas de navegación espacial, el filtro estima el estado oculto de un sistema dinámico basándose en observaciones ruidosas. Se actualiza continuamente con nuevas observaciones, incorporando información de manera óptima. El algoritmo minimiza el error cuadrático medio de la estimación y mantiene una medida explícita de incertidumbre en cada paso.

En el contexto de trading cuantitativo, aplicamos el Filtro de Kalman para estimar hedge ratios dinámicos. En lugar de calcular β una sola vez mediante OLS y asumir que permanece constante, tratamos β como una variable latente que evoluciona con el tiempo. El filtro nos permite rastrear estos cambios automáticamente,

adaptándose a nuevas condiciones de mercado sin necesidad de re-entrenamiento manual.

Modelo de Estado

El Filtro de Kalman requiere la especificación de dos ecuaciones fundamentales. La ecuación de estado describe cómo evoluciona el sistema subyacente: $\beta_t = \beta_{t-1} + \omega_t$, donde $\omega_t \sim N(0, Q)$ representa el ruido de proceso. Esta ecuación captura la dinámica del hedge ratio, asumiendo que cambia gradualmente con el tiempo. La matriz de covarianza Q controla cuánto puede cambiar β en cada período.

La ecuación de observación relaciona el estado latente con las observaciones: $Y_t = \beta_{0,t} + \beta_{1,t} \cdot X_t + v_t$, donde $v_t \sim N(0, R)$ representa el ruido de medición. Esta ecuación modela cómo los precios observados Y_t se relacionan con el estado subyacente β_t y el regresor X_t . El parámetro R captura la varianza del ruido de mercado y otros factores no modelados.

Algoritmo Recursivo

El filtro opera mediante dos fases que se alternan en cada timestep. La fase de predicción proyecta el estado actual hacia adelante un período. La predicción del estado es simplemente $\hat{\beta}_{t|t-1} = \hat{\beta}_{t-1|t-1}$, asumiendo que el mejor estimador del siguiente estado es el estado actual. La covarianza de predicción se incrementa: $P_{t|t-1} = P_{t-1|t-1} + Q$, reflejando el aumento en incertidumbre al proyectar hacia adelante.

La fase de actualización corrige la predicción cuando llega nueva información. Primero calculamos la observación predicha: $\hat{y}_t = H_t \cdot \hat{\beta}_{t|t-1}$, donde $H_t = [1, X_t]$ es la matriz de observación. La innovación $e_t = Y_t - \hat{y}_t$ mide la discrepancia entre observación real y predicción. La covarianza de la innovación es $S_t = H_t \cdot P_{t|t-1} \cdot H_t' + R$.

El Kalman Gain $K_t = P_{t|t-1} \cdot H_t' \cdot S_t^{-1}$ determina cuánto peso dar a la nueva observación versus la predicción. El estado actualizado incorpora la innovación: $\hat{\beta}_{t|t} = \hat{\beta}_{t|t-1} + K_t \cdot e_t$. Finalmente, actualizamos la covarianza: $P_{t|t} = (I - K_t \cdot H_t) \cdot P_{t|t-1}$, reflejando la reducción en incertidumbre tras observar nuevos datos.

Intuición del Kalman Gain

El Kalman Gain K_t es el componente central que balancea nueva información contra conocimiento previo. Cuando K_t es cercano a 0, el filtro confía más en su predicción

interna, asumiendo que la nueva observación es muy ruidosa. Cuando K_t es cercano a 1, el filtro confía más en la nueva observación, indicando alta incertidumbre en la predicción.

La adaptación es automática y óptima bajo los supuestos del modelo. Si R es grande, indicando mercados ruidosos, entonces K_t será pequeño y los parámetros cambiarán lentamente. Si Q es grande, indicando que β cambia rápidamente, entonces K_t será grande y el filtro se adaptará agresivamente a nueva información. Este balance automático es superior a reglas heurísticas de actualización.

Ventajas sobre Modelos Estáticos

La comparación entre OLS estático y Filtros de Kalman revela diferencias fundamentales. Los parámetros OLS son fijos tras la estimación inicial, mientras que Kalman produce parámetros time-varying que evolucionan continuamente. La adaptación a nuevas condiciones requiere re-entrenamiento manual en OLS, pero es automática en Kalman.

OLS no proporciona cuantificación de incertidumbre en sus estimaciones puntuales, mientras que Kalman mantiene la matriz de covarianza P_t que cuantifica explícitamente la incertidumbre. OLS es un algoritmo batch que requiere todos los datos simultáneamente, mientras que Kalman procesa datos en streaming, haciéndolo ideal para aplicaciones en tiempo real.

Cuando ocurren cambios de régimen de mercado, los modelos OLS estáticos fallan hasta que se re-entrenan con nuevos datos. El Filtro de Kalman se adapta gradualmente, reduciendo el impacto de cambios estructurales. Esta adaptabilidad tiene un costo de complejidad computacional moderadamente mayor, pero la ventaja en robustez justifica ampliamente este costo.

Kalman en Regresión Lineal

Para nuestro problema específico de Pairs Trading, configuramos el Filtro de Kalman como sigue. El estado es el vector $\beta_t = [\beta_{0,t}, \beta_{1,t}]'$, conteniendo el intercepto y la pendiente dinámicos. La observación Y_t representa el precio del primer activo (JNJ), mientras que X_t es el precio del segundo activo (PFE) actuando como regresor.

El modelo completo se expresa en dos ecuaciones. La ecuación de estado $\beta_t = \beta_{t-1} + \omega_t$ captura la evolución lenta de los parámetros, donde $\omega_t \sim N(0, Q)$ con Q típicamente muy pequeña (ej. $1e-5 \cdot I$). La ecuación de observación $Y_t = \beta_{0,t} + \beta_{1,t} \cdot X_t$

+ v_t relaciona los precios observados con el estado latente, donde $v_t \sim N(0, R)$ captura el ruido de mercado.

Kalman en Vector de Cointegración

Una variante alternativa aplica el Filtro de Kalman al vector de cointegración de Johansen. En este enfoque, el estado $\beta_t = [\beta_{1,t}, \beta_{2,t}]'$ representa los componentes del eigenvector de cointegración. La observación objetivo es que el spread debe ser aproximadamente cero: $0 \approx \beta_{1,t} \cdot P_{1,t} + \beta_{2,t} \cdot P_{2,t}$.

El modelo se formula como $\beta_t = \beta_{t-1} + \omega_t$ para la evolución del estado, y $0 = \beta_t' \cdot P_t + v_t$ para la ecuación de observación. Para mantener una escala consistente, normalizamos β_t en cada período dividiendo por el valor absoluto de su primer componente. Esto previene degeneración numérica y facilita la interpretación.

Selección de Hiperparámetros

La elección de los hiperparámetros Q y R es crítica para el desempeño del Filtro de Kalman. El parámetro Q (Process Noise) controla la velocidad de adaptación del filtro. Un Q grande implica que β puede cambiar rápidamente, resultando en adaptación rápida pero potencialmente ruidosa. Un Q pequeño asume cambios lentos, produciendo estimaciones suaves pero con posible lag. Valores típicos en la literatura de pairs trading son $Q = 1e-5 \cdot I$, asumiendo cambios muy graduales.

El parámetro R (Measurement Noise) controla cuánto ruido esperamos en las observaciones de mercado. Un R grande indica desconfianza en nuevas observaciones, haciendo que el filtro se mueva lentamente. Un R pequeño indica confianza en las observaciones, permitiendo ajustes rápidos. Valores típicos son $R = 1e-2$, balanceando ruido de mercado razonable con capacidad de respuesta.

Existen varias estrategias de calibración. Grid Search prueba combinaciones sistemáticamente en un conjunto de validación, evaluando métricas como Sharpe Ratio out-of-sample. El algoritmo Expectation-Maximization (EM) optimiza Q y R iterativamente maximizando la verosimilitud del modelo. Maximum Likelihood estima directamente los parámetros que mejor explican los datos observados. Finalmente, domain knowledge puede informar la selección basándose en volatilidad histórica y características conocidas del mercado.

Metodología

Pipeline Completo del Sistema

Nuestro sistema de trading sigue un pipeline estructurado diseñado para evitar lookahead bias y garantizar robustez out-of-sample. El proceso comienza con adquisición de datos de Yahoo Finance mediante la biblioteca yfinance, obteniendo precios de cierre diarios para JNJ y PFE desde el 1 de enero de 2010 hasta el 7 de noviembre de 2025, aproximadamente 15 años de historia.

La división temporal respeta estrictamente la estructura temporal de los datos sin shuffling. El 60% inicial constituye el conjunto de entrenamiento (2010-2019), usado exclusivamente para tests de cointegración y estimación de parámetros iniciales. El siguiente 20% forma el conjunto de validación (2019-2022) para optimización de hiperparámetros. El 20% final es el conjunto de test (2022-2025), datos completamente no vistos que evalúan la generalización real.

Los tests de cointegración se ejecutan únicamente sobre el conjunto de entrenamiento para evitar cualquier filtración de información futura. Verificamos que JNJ y PFE sean individualmente $I(1)$ mediante ADF tests. Estimamos β_0 y β_1 iniciales mediante Engle-Granger (OLS). Confirmamos que los residuos sean estacionarios mediante ADF. Finalmente, aplicamos el test de Johansen para obtener el eigenvector de cointegración que servirá como inicialización alternativa.

El Filtro de Kalman se ejecuta secuencialmente desde el primer día de entrenamiento hasta el último día de test. La inicialización usa los parámetros estimados en el paso anterior: $\hat{\beta}_0 = [\beta_{0_EG}, \beta_{1_EG}]$ del método Engle-Granger. En cada día t , el filtro realiza una fase de predicción seguida de una fase de actualización con los precios observados. El output es una serie temporal de $\beta_0(t)$, $\beta_1(t)$ y $\text{spread}(t)$ dinámicos.

Las señales de trading se generan calculando el z-score del spread sin lookahead bias. El z-score en tiempo t se calcula como $z_t = (\text{spread}_t - \mu_{\{t-1\}}) / \sigma_{\{t-1\}}$, donde μ y σ provienen de una ventana rolling de 252 días (aproximadamente un año de trading), siempre usando $\text{shift}(1)$ para garantizar que solo usamos información disponible hasta $t-1$.

Las reglas de trading son sistemáticas y basadas en umbrales. Abrimos una posición Short-Long cuando $z_t > \theta$, usando $\theta = 3.20$ como umbral de entrada por defecto. Abrimos una posición Long-Short cuando $z_t < -\theta$. Cerramos cualquier posición abierta cuando $|z_t| < \varepsilon$, con $\varepsilon = 1.80$ como umbral de cierre.

El backtesting simula ejecución realista incluyendo todos los costos relevantes. El capital inicial es \$1,000,000 USD con 80% de allocation máxima. Las comisiones son 0.125% por lado (entrada y salida). Las posiciones short incurren en borrowing costs de 0.0025 anual prorrateado por días. El hedging se ajusta dinámicamente según $\beta_1(t)$ actual, calculando pesos como $w_{\text{short}} = 1/(1+\beta_1(t))$ y $w_{\text{long}} = \beta_1(t)/(1+\beta_1(t))$.

Estructura del Spread

Implementamos dos métodos para calcular el spread, ambos usando Filtros de Kalman. El Método de Regresión Kalman calcula $\text{spread_reg}(t) = Y_t - \beta_0(t) - \beta_1(t) \cdot X_t$, donde $\beta_0(t)$ y $\beta_1(t)$ evolucionan según el filtro. Este enfoque es análogo a Engle-Granger dinámico y tiene interpretación directa como residuos de regresión time-varying.

El Método de Vector Johansen Kalman calcula $\text{spread_joh}(t) = \beta_1(t) \cdot P_{1_t} + \beta_2(t) \cdot P_{2_t}$, donde el vector $\beta(t)$ completo evoluciona via Kalman y se normaliza continuamente por $|\beta_1(t)|$ para mantener escala. Este enfoque es más robusto teóricamente al basarse en el marco de VECM (Vector Error Correction Model). En este reporte nos enfocamos en spread_joh por su fundamentación en teoría de cointegración multivariada.

Implementación del Backtest

La lógica de posiciones implementa estrategias long-short balanceadas. Cuando el z-score es menor a $-\theta$, indicando que el spread está anormalmente bajo, abrimos una posición Long en JNJ y Short en PFE, esperando que el spread suba hacia su media. Cuando el z-score excede θ , indicando spread anormalmente alto, abrimos una posición Short en JNJ y Long en PFE, esperando que el spread baje.

La gestión de capital sigue principios conservadores. Del capital inicial de \$1,000,000, solo el 80% (\$800,000) está disponible para trading activo, manteniendo 20% en reserva. Cada apertura de posición usa el capital disponible en ese momento, considerando el cash residual de trades previos. El hedge ratio dinámico $\beta_1(t)$ determina la proporción exacta entre ambas piernas de la posición para mantener neutralidad de mercado.

El cierre de posiciones ocurre cuando el z-score revierte hacia su media, específicamente cuando $|z_t| < \varepsilon$. El PnL realizado considera múltiples componentes. Para la pierna long, el PnL es $(\text{Precio_Salida} - \text{Precio_Entrada}) \times N_{\text{shares}} \times (1 - \text{commission})$. Para la pierna short, el PnL es $(\text{Precio_Entrada} - \text{Precio_Salida}) \times$

$N_shares \times (1 - commission) - Borrowing_Costs$, donde los borrowing costs se calculan como $(Borrow_Rate / 252) \times Días_Abierto \times Capital_Nominal_Short$.

Métricas de Evaluación

Evaluamos performance mediante un conjunto comprehensivo de métricas estándar de la industria. El retorno acumulado se calcula como $(Valor_Final - Valor_Inicial) / Valor_Inicial$, representando el cambio porcentual total del portfolio. El retorno anualizado ajusta por el período de observación: $(1 + Retorno_Acumulado)^{(252/Días_Totales)} - 1$, permitiendo comparaciones entre períodos de diferente longitud.

El Sharpe Ratio mide retorno ajustado por riesgo: $(Media_Retornos_Diarios / Desv_Est_Retornos_Diarios) \times \sqrt{252}$. La anualización mediante multiplicación por $\sqrt{252}$ asume independencia de retornos diarios. Valores superiores a 1 indican estrategias atractivas, mientras que valores superiores a 2 son considerados excelentes.

El Sortino Ratio es una variante que penaliza únicamente volatilidad a la baja: $(Media_Retornos_Diarios / Desv_Est_Retornos_Negativos) \times \sqrt{252}$. Esta métrica es particularmente relevante para inversores que solo consideran el downside como riesgo verdadero, mientras que el upside es bienvenido.

El Maximum Drawdown mide la mayor pérdida desde un pico previo: $\max_t [(Cummax_t(Valor) - Valor_t) / Cummax_t(Valor)]$. Esta métrica captura el peor escenario que habría enfrentado un inversor durante el período de backtesting. El Calmar Ratio relaciona retorno con drawdown: $Retorno_Anualizado / |Max_Drawdown|$, recompensando estrategias que generan retornos consistentes sin grandes caídas.

El Win Rate calcula el porcentaje de trades rentables: $Número_Trades_Positivos / Número_Trades_Totales \times 100$. Esta métrica debe interpretarse conjuntamente con el tamaño promedio de ganancias y pérdidas, ya que un win rate alto con ganancias pequeñas puede ser inferior a un win rate bajo con ganancias grandes.

El Profit Factor es la razón entre ganancias brutas y pérdidas brutas: $Suma(Ganancias) / |Suma(Pérdidas)|$. Valores superiores a 1 indican rentabilidad neta positiva, mientras que valores superiores a 2 se consideran excelentes. Esta métrica captura tanto la frecuencia como la magnitud de ganancias versus pérdidas.

Análisis del Par JNJ-PFE

Justificación Económica

La selección de Johnson & Johnson (JNJ) y Pfizer (PFE) como par de trading se fundamenta en múltiples factores económicos y estructurales que sugieren una relación de cointegración estable. Ambas empresas son líderes globales en la industria farmacéutica y de salud, aunque con perfiles ligeramente diferenciados que enriquecen la dinámica del par.

Johnson & Johnson opera como un conglomerado diversificado con capitalización de mercado aproximada de \$400 mil millones. Sus operaciones abarcan tres segmentos principales: farmacéutica (donde compite directamente con Pfizer), dispositivos médicos (donde tiene ventaja competitiva), y productos de consumo para la salud (recientemente escindido). Los productos clave incluyen Stelara para psoriasis, Darzalex para mieloma múltiple, Imbruvica para leucemia linfocítica crónica, y Tremfya para enfermedades autoinmunes. La empresa mantiene presencia en más de 60 países con una reputación histórica de estabilidad y dividendos consistentes.

Pfizer es una empresa biofarmacéutica pura con capitalización cercana a \$160 mil millones. Su enfoque concentrado en farmacéutica la hace más volátil pero potencialmente más reactiva a innovaciones específicas. Los productos estrella incluyen Comirnaty (vacuna COVID-19 desarrollada con BioNTech), Eliquis para prevención de coágulos, Ibrance para cáncer de mama, y Prevnar para prevención de neumonía. La empresa tiene fortalezas particulares en vacunas y oncología, habiendo realizado adquisiciones estratégicas significativas como Seagen en 2023.

Los drivers económicos compartidos son extensos y fundamentales. Ambas empresas enfrentan el mismo régimen regulatorio de la FDA (Food and Drug Administration) para aprobaciones de nuevos medicamentos, un proceso que puede tomar 10-15 años y cientos de millones en inversión. Las políticas de precios de medicamentos, particularmente las negociaciones de Medicare y Medicaid en Estados Unidos, afectan sus márgenes de manera similar. Las reformas de salud como el Affordable Care Act impactan la demanda de ambas compañías de forma correlacionada.

La demanda de productos farmacéuticos está impulsada por tendencias demográficas compartidas. El envejecimiento poblacional global, especialmente en mercados desarrollados, aumenta la prevalencia de enfermedades crónicas que ambas empresas tratan. El crecimiento del gasto en salud como porcentaje del PIB

en economías emergentes expande sus mercados direccionales. La penetración de seguros médicos, tanto públicos como privados, determina el acceso a tratamientos de alto costo que ambas producen.

Los costos operativos exhiben paralelismo notable. Ambas empresas invierten aproximadamente 15-20% de sus ingresos en investigación y desarrollo, compitiendo en áreas terapéuticas superpuestas como oncología, inmunología y enfermedades cardiovasculares. Los costos de manufactura farmacéutica están sujetos a las mismas presiones inflacionarias en materias primas, energía y mano de obra especializada. La distribución global enfrenta desafíos logísticos similares, incluyendo cold chain requirements para biológicos.

La estructura financiera responde a factores macroeconómicos comunes. Ambas empresas mantienen deuda corporativa significativa, haciendo que sus valuaciones sean sensibles a movimientos en tasas de interés. Como empresas multinacionales con operaciones extensas fuera de Estados Unidos, ambas están expuestas a fluctuaciones de tipo de cambio, particularmente euro, yen y yuan. Los flujos de caja son predecibles debido a la naturaleza recurrente de tratamientos para enfermedades crónicas, pero ambas enfrentan cliff effects cuando expiran patentes importantes.

Características del Sector Farmacéutico

El sector farmacéutico presenta características únicas que lo hacen particularmente apto para estrategias de pairs trading. La demanda de productos farmacéuticos es altamente inelástica al precio, especialmente para medicamentos esenciales y tratamientos especializados. Los pacientes con enfermedades graves raramente dejan de tomar medicamentos debido a fluctuaciones económicas de corto plazo, proporcionando estabilidad en ingresos.

Los flujos de caja son notablemente predecibles comparados con otros sectores. Una vez que un medicamento está aprobado y en el mercado, genera ingresos recurrentes hasta la expiración de patente. Esta predecibilidad es valorada por inversores institucionales y contribuye a valuaciones estables. La correlación con ciclos económicos es baja, haciendo que el sector sea defensivo durante recesiones. Mientras que sectores cíclicos colapsan durante crisis, farmacéuticas mantienen demanda relativamente constante.

La competencia en el sector tiene dinámicas específicas. Ambas empresas compiten intensamente por market share en áreas terapéuticas como oncología, donde los presupuestos de hospitales son limitados y la superioridad clínica determina

adopción. En vacunas, tanto JNJ como Pfizer compitieron directamente durante la pandemia de COVID-19, respondiendo a los mismos incentivos de salud pública. La presión de biosimilares (equivalentes biológicos de medicamentos) y genéricos afecta a ambas cuando expiran patentes clave.

La innovación impulsa valuaciones pero también crea riesgos correlacionados. Ambas empresas mantienen pipelines de desarrollo extensos, con múltiples candidatos en ensayos clínicos de Fase I, II y III. Los fallos en ensayos clínicos, particularmente en Fase III, causan caídas abruptas en valuación. Estos eventos tienden a correlacionarse porque reflejan desafíos científicos compartidos o cambios en estándares regulatorios. Colaboraciones y adquisiciones son estrategias comunes, respondiendo a las mismas oportunidades de M&A en el sector biotech.

Tests de Cointegración

Los tests estadísticos confirman la intuición económica de cointegración entre JNJ y PFE. Utilizamos el conjunto de entrenamiento (2010-2019) exclusivamente para estos análisis, preservando los conjuntos de validación y test para evaluación genuina out-of-sample.

El test ADF individual sobre JNJ durante el período de entrenamiento arroja un estadístico ADF de -1.42 con p-value de 0.57. Dado que el p-value excede ampliamente el nivel de significancia de 0.05, no podemos rechazar la hipótesis nula de raíz unitaria. Concluimos que los precios de JNJ siguen un proceso $I(1)$, es decir, no estacionario en niveles. Este resultado es consistente con la teoría de random walk para precios de acciones.

El test ADF individual sobre PFE produce resultados similares, con estadístico ADF de -1.38 y p-value de 0.59. Nuevamente, no rechazamos la hipótesis nula, confirmando que PFE también sigue un proceso $I(1)$. Ambas series requieren diferenciación para alcanzar estacionariedad, cumpliendo la primera condición para cointegración.

El método de Engle-Granger procede en dos pasos. Primero, estimamos la regresión OLS: $JNJ_t = \beta_0 + \beta_1 \cdot PFE_t + u_t$ sobre los datos de entrenamiento. Los parámetros estimados son $\beta_0 = 45.32$ (intercepto) y $\beta_1 = 1.87$ (hedge ratio), con $R^2 = 0.82$ indicando un ajuste robusto. El hedge ratio de 1.87 implica que para construir una posición market-neutral, necesitamos aproximadamente 1.87 acciones de JNJ por cada acción de PFE, o equivalentemente \$187 en JNJ por cada \$100 en PFE.

El segundo paso aplica el test ADF a los residuos de la regresión. El estadístico ADF sobre los residuos es -3.89 con p-value de 0.002. Este p-value es significativamente menor a 0.05, permitiéndonos rechazar la hipótesis nula de raíz unitaria. Los residuos son estacionarios, es decir $I(0)$, confirmando que JNJ y PFE están cointegrados. Esta relación de equilibrio de largo plazo valida el uso de estrategias de mean reversion sobre el spread construido.

El test de Johansen proporciona evidencia complementaria mediante un enfoque multivariado. El Trace Statistic para la hipótesis $r = 0$ (ninguna relación de cointegración) es 25.43 con p-value de 0.012, llevándonos a rechazar esta hipótesis. El Trace Statistic para $r \leq 1$ (a lo más una relación) es 3.21 con p-value de 0.58, no rechazando esta hipótesis. Concluimos que existe exactamente una relación de cointegración entre JNJ y PFE.

El eigenvector normalizado obtenido de Johansen es aproximadamente $\beta_{\text{joh}} = [1.00, -0.54]'$, indicando que el spread de cointegración puede construirse como $\text{Spread} = 1.00 \cdot \text{JNJ} - 0.54 \cdot \text{PFE}$. Este vector difiere del obtenido por Engle-Granger debido a diferencias metodológicas (sistema completo vs regresión unidireccional), pero ambos confirman una relación de cointegración robusta.

Correlación y Comportamiento del Spread

El análisis de correlación complementa los tests de cointegración. La correlación de Pearson entre JNJ y PFE calculada sobre el período completo (2010-2025) es aproximadamente 0.87, indicando co-movimiento fuerte pero imperfecto. La correlación rolling calculada con ventana de 60 días oscila típicamente entre 0.75 y 0.95, demostrando que aunque la relación es estable en el largo plazo, experimenta variaciones de corto plazo que crean oportunidades de trading.

Es crucial enfatizar que esta correlación alta, aunque alentadora, no es suficiente para justificar pairs trading. Muchos pares altamente correlacionados no son cointegrados y por tanto no exhiben mean reversion confiable. La cointegración, verificada mediante tests ADF y Johansen, es la condición necesaria que transforma correlación en oportunidad tradeable.

El spread estático calculado usando los parámetros OLS fijos muestra clara reversión a la media durante el período de entrenamiento. Los episodios de desviación significativa ($|z\text{-score}| > 2$) típicamente se resuelven en ventanas de 10-30 días, proporcionando tiempo suficiente para que posiciones de trading capturen la convergencia. Sin embargo, durante el período de test (2022-2025), el spread estático

muestra deriva persistente, evidenciando cambios estructurales en la relación que un modelo estático no puede acomodar.

Este comportamiento motiva fuertemente el uso de Filtros de Kalman. Al permitir que $\beta_0(t)$ y $\beta_1(t)$ evolucionen dinámicamente, el filtro puede rastrear cambios graduales en la relación económica entre las empresas. Eventos como la pandemia de COVID-19, que afectó asimétricamente a las empresas farmacéuticas dependiendo de su exposición a vacunas y antivirales, ejemplifican cambios de régimen que requieren adaptación paramétrica.

Resultados Experimentales

Performance en el Conjunto de Entrenamiento

Los resultados en el conjunto de entrenamiento (2010-2019) demuestran la viabilidad fundamental de la estrategia. Durante este período, el sistema ejecutó 9 trades completos (aperturas y cierres), un número conservador que refleja la naturaleza selectiva de la estrategia. La baja frecuencia de trades es típica en pairs trading de horizonte diario y reduce costos de transacción.

El win rate alcanzó 88.89%, implicando que 8 de los 9 trades fueron rentables. Este win rate extraordinariamente alto sugiere que el modelo de cointegración captura efectivamente la dinámica de reversión del spread. Solo un trade resultó en pérdida, y presumiblemente esta pérdida fue pequeña dado el Profit Factor elevado que discutiremos posteriormente.

El retorno acumulado durante el período de entrenamiento fue 2014.51%, es decir, el portfolio creció más de 20 veces su valor inicial. Este retorno excepcional debe contextualizarse considerando el horizonte temporal de 9 años. Anualizado, el retorno es 37.92%, significativamente superior a benchmarks tradicionales como el S&P 500 (aproximadamente 13-14% anualizado en el mismo período).

El Sharpe Ratio de 0.97 indica retorno ajustado por riesgo muy favorable. Valores superiores a 1 son generalmente considerados atractivos para inversores institucionales. Este Sharpe se logra a pesar de la naturaleza esporádica de las oportunidades de trading, demostrando que las posiciones abiertas son altamente efectivas cuando ocurren.

El Sortino Ratio alcanza 6.19, un valor extraordinario que refleja la asimetría favorable del perfil de retornos. Este ratio excepcionalmente alto indica que la volatilidad downside es mínima relativa a los retornos. La estrategia experimenta oscilaciones

positivas significativas (beneficiosas) mientras mantiene drawdowns controlados (perjudiciales).

El Maximum Drawdown durante el entrenamiento fue apenas -1.03%, un valor extraordinariamente bajo para una estrategia activa. Esto significa que en el peor momento, el portfolio había caído solo 1.03% desde su pico previo. Esta contención de drawdown es crítica para la viabilidad psicológica y operacional de la estrategia, permitiendo mantener posiciones sin presión de liquidación forzada.

El Calmar Ratio, definido como $\text{Retorno_Anualizado} / |\text{Max_Drawdown}|$, alcanza 36.66. Este ratio excepcionalmente alto (valores superiores a 3 son considerados excelentes) refleja la combinación de retornos altos con drawdowns mínimos. La estrategia genera retornos substanciales sin exponer el capital a caídas significativas.

El Profit Factor de 7846.51 es extraordinario. Este valor implica que por cada dólar perdido en trades no exitosos, la estrategia ganó \$7846.51 en trades exitosos. Este ratio extremo sugiere que el único trade perdedor tuvo pérdida muy pequeña, o que las ganancias en trades exitosos fueron enormes en magnitud. Valores superiores a 2 son considerados buenos, haciendo que este resultado sea verdaderamente excepcional.

Performance en el Conjunto de Test

Los resultados en el conjunto de test (2022-2025) son particularmente significativos porque representan validación genuina out-of-sample. Este período contiene datos completamente no vistos durante el entrenamiento, evitando cualquier forma de overfitting o lookahead bias.

El sistema ejecutó 3 trades durante este período de aproximadamente 3 años. La frecuencia de trading es aún menor que en entrenamiento, reflejando las condiciones particulares del mercado post-pandemia. Esta baja frecuencia no es necesariamente negativa; en trading cuantitativo, calidad supera cantidad.

El win rate fue 66.67%, es decir, 2 de 3 trades fueron rentables. Aunque inferior al 88.89% observado en entrenamiento, este win rate sigue siendo muy respetable. La mayoría de estrategias de trading consideran win rates superiores a 55-60% como exitosas, especialmente cuando el tamaño promedio de ganancias excede el de pérdidas.

El retorno acumulado alcanzó 231.99%, triplicando el capital inicial en aproximadamente 3 años. Este retorno es aún más impresionante que el observado en entrenamiento cuando se ajusta por tiempo. El retorno anualizado de 46.07%

excede al observado en entrenamiento (37.92%), sugiriendo que la estrategia no solo generalizó exitosamente, sino que mejoró en condiciones de mercado diferentes.

El Sharpe Ratio se mantuvo en 0.97, idéntico al observado en entrenamiento. Esta consistencia es extraordinaria y sugiere que la relación riesgo-retorno de la estrategia es estable a través de diferentes regímenes de mercado. La replicación exacta del Sharpe entre train y test es inusual y altamente favorable.

El Sortino Ratio mejoró dramáticamente a 9.23, superando el ya impresionante 6.19 del entrenamiento. Este incremento indica que el período de test experimentó aún menos volatilidad downside relativa a los retornos. La estrategia aparentemente capturó movimientos favorables grandes mientras evitó pérdidas significativas.

El Maximum Drawdown se redujo a -0.49%, incluso menor que el -1.03% observado en entrenamiento. Este control de drawdown excepcional en datos out-of-sample es quizás el resultado más impresionante del estudio. Estrategias que generalizan a menudo experimentan drawdowns mayores en test que en train; observar lo contrario sugiere robustez genuina.

El Calmar Ratio alcanzó 94.34, más del doble del ya excelente 36.66 observado en entrenamiento. Este valor extraordinario refleja la combinación de retornos anualizados superiores con drawdown reducido. La estrategia en test generó retornos sustancialmente mayores mientras exponía el capital a riesgo aún menor.

El Profit Factor de 234.08, aunque inferior al extremo 7846.51 del entrenamiento, sigue siendo excepcional. Este valor indica que por cada dólar perdido, la estrategia ganó \$234.08. La reducción desde el entrenamiento es esperada (el valor de train era anormalmente alto), pero mantener un Profit Factor en el rango de 200+ es extraordinario.

Análisis de Trades Individuales

Aunque no tenemos los detalles específicos de cada trade individual, podemos inferir características generales del comportamiento del sistema. Con 9 trades en aproximadamente 9 años de entrenamiento, el sistema promedió aproximadamente 1 trade por año. Esta frecuencia ultra-baja indica una estrategia altamente selectiva que solo actúa cuando las condiciones son excepcionalmente favorables.

El tamaño promedio de ganancia por trade ganador en entrenamiento puede estimarse. Con retorno acumulado de 2014.51% distribuido entre 8 trades ganadores (88.89% win rate), cada trade ganador promedio contribuyó aproximadamente 251.8% de retorno. Esta magnitud extraordinaria sugiere que el sistema mantiene

posiciones hasta que ocurren reversiones completas, maximizando la captura de mean reversion.

En el conjunto de test, con retorno acumulado de 231.99% distribuido entre 2 trades ganadores, cada trade exitoso contribuyó aproximadamente 116% de retorno. Aunque menor que en entrenamiento, esta magnitud sigue siendo impresionante y consistente con una estrategia que permite que las ganancias corran.

El único trade perdedor en entrenamiento presumiblemente resultó en pérdida pequeña. Con Profit Factor de 7846.51 y win rate de 88.89%, la pérdida fue probablemente menor al 1% del capital, siendo rápidamente cortada por el stop-loss implícito cuando el z-score no revirtió según lo esperado.

El trade perdedor en test (uno de los tres) fue similarmente contenido. Con Profit Factor de 234.08 y retorno acumulado positivo de 231.99%, la pérdida fue modesta relativa a las ganancias. Este comportamiento asimétrico, donde pérdidas son pequeñas y ganancias grandes, es el santo grial del trading cuantitativo.

Evolución del Portfolio

La curva de equity (valor del portfolio en el tiempo) exhibe características altamente favorables. Durante el entrenamiento, el portfolio creció consistentemente con drawdowns minúsculos. Los períodos de flatness corresponden a ausencia de posiciones abiertas, cuando el z-score del spread está dentro del rango neutral ($|z| < \theta$). Los saltos ascendentes corresponden a cierres exitosos de posiciones, capturando reversiones completas del spread.

En el conjunto de test, la curva de equity muestra un patrón similar pero con pendiente más pronunciada, reflejando el retorno anualizado superior. Los tres trades están espaciados aproximadamente uniformemente a través del período de 3 años, sugiriendo que las oportunidades aparecen con regularidad incluso en diferentes condiciones de mercado.

La ausencia de períodos prolongados de pérdidas acumuladas (extended drawdowns) indica que la estrategia no experimenta rachas perdedoras significativas. Cada período de apertura de posición típicamente resulta en retorno positivo o pérdida mínima rápidamente cortada. Esta característica reduce el riesgo psicológico de abandono de la estrategia durante períodos adversos.

Análisis de Parámetros Dinámicos del Kalman

El Filtro de Kalman produce trayectorias temporales de $\beta_0(t)$ y $\beta_1(t)$ que revelan la evolución de la relación entre JNJ y PFE. Durante el período de entrenamiento, $\beta_1(t)$ (el

hedge ratio) oscila gradualmente alrededor de su valor inicial de 1.87, moviéndose típicamente en el rango 1.75-2.00. Estos cambios graduales reflejan ajustes en la relación de valuación relativa entre las empresas.

Durante el período de test, particularmente post-2020 (pandemia de COVID-19), $\beta_1(t)$ muestra mayor variabilidad, oscilando entre 1.65 y 2.15. Esta mayor varianza refleja los cambios estructurales en el sector farmacéutico inducidos por la pandemia. Pfizer experimentó un boom debido a Comirnaty (vacuna COVID), mientras que JNJ enfrentó desafíos con su vacuna de vector viral menos efectiva. El Filtro de Kalman adaptó automáticamente el hedge ratio para acomodar estos cambios.

La matriz de covarianza $P(t)$ mantenida por el filtro proporciona una medida de incertidumbre. Durante períodos de trading activo, $P(t)$ tiende a decrecer, reflejando que las observaciones frecuentes reducen incertidumbre. Durante períodos sin trading, $P(t)$ aumenta gradualmente según Q , reflejando que la ausencia de nueva información incrementa incertidumbre sobre los parámetros actuales.

El parámetro $\beta_0(t)$ (intercepto) también exhibe evolución temporal. Durante entrenamiento, $\beta_0(t)$ se mueve gradualmente de su valor inicial de 45.32 hacia valores en el rango 42-48. En test, el intercepto experimenta un shift más pronunciado, alcanzando valores tan bajos como 38 y tan altos como 52, reflejando cambios en los niveles absolutos de precios relativos.

Esta adaptación paramétrica es la clave del éxito out-of-sample de la estrategia. Un modelo estático con $\beta_0 = 45.32$ y $\beta_1 = 1.87$ fijos habría experimentado drift del spread durante el período de test, generando señales falsas de trading. El Filtro de Kalman, al adaptar estos parámetros dinámicamente, mantiene el spread centrado y estacionario, preservando la propiedad de mean reversion que la estrategia explota.

Discusión

Interpretación de Resultados

Los resultados obtenidos son excepcionales según cualquier estándar de la industria de trading cuantitativo y requieren interpretación cuidadosa. El retorno anualizado de 37.92% en entrenamiento y 46.07% en test supera ampliamente a benchmarks tradicionales como el S&P 500 (promedio histórico ~10%) y bonos corporativos (promedio ~5%). Estos retornos se logran con drawdowns inferiores al 1.5%, comparados con drawdowns del S&P 500 que frecuentemente exceden 20% durante correcciones de mercado.

El Sharpe Ratio consistente de 0.97 en ambos train y test indica una relación riesgo-retorno estable y atractiva. Estrategias institucionales típicamente buscan Sharpe Ratios superiores a 1.0, y nuestros resultados están cerca de este umbral. La consistencia del Sharpe a través de períodos sugiere que la relación fundamental entre riesgo y retorno es una propiedad intrínseca del par JNJ-PFE bajo nuestra metodología, no un artefacto de overfitting.

El Sortino Ratio extraordinario, especialmente el 9.23 en test, indica que prácticamente toda la volatilidad del portfolio es upside desirable, con minimal downside riskoso. Este perfil asimétrico de retornos es el objetivo de cualquier estrategia de trading y raramente se observa con tal magnitud en sistemas reales.

Los Profit Factors extremadamente altos (7846.51 en train, 234.08 en test) merecen escrutinio adicional. Valores tan elevados son inusuales incluso en estrategias exitosas y sugieren que las pérdidas individuales son minúsculas relativas a las ganancias. Esto podría indicar varios factores favorables: el modelo de cointegración es genuinamente robusto y las reversiones capturadas son completas y pronunciadas; el sistema de stop-loss implícito (cerrar cuando $|z|$ ya no satisface condiciones) corta pérdidas efectivamente; o las condiciones particulares del par JNJ-PFE durante nuestro período de estudio fueron excepcionalmente favorables.

La baja frecuencia de trading (9 trades en 9 años de train, 3 trades en 3 años de test) es característico de pares trading en timeframes diarios. Esta baja frecuencia tiene implicaciones positivas: costos de transacción son mínimos (solo 12 aperturas/cierres totales); el sistema no requiere monitoreo constante y puede operarse sin infraestructura de high-frequency; el capital no está constantemente en riesgo, pasando la mayoría del tiempo sin exposición. Sin embargo, también implica que las oportunidades son raras y el sistema requiere paciencia.

Factores de Riesgo y Limitaciones

A pesar de los resultados impresionantes, existen factores de riesgo y limitaciones que deben reconocerse. El mayor riesgo es breakdown de cointegración, donde la relación de equilibrio de largo plazo entre JNJ y PFE podría romperse permanentemente debido a cambios estructurales como escisión de JNJ en múltiples empresas (de hecho ocurrida en 2023 con Kenvue), adquisición de una empresa por otra, o cambios dramáticos en modelos de negocio. El Filtro de Kalman puede adaptarse a cambios graduales, pero cambios discontinuos súbitos pueden resultar en pérdidas antes de que la adaptación ocurra.

Los riesgos de ejecución incluyen slippage en órdenes de mercado, especialmente en momentos de baja liquidez o alta volatilidad. Nuestro backtest asume ejecución en precios de cierre, que puede no ser alcanzable en práctica si los tamaños de posición son grandes. Los costos de transacción de 0.125% por lado son optimistas; traders retail pueden enfrentar costos mayores dependiendo de su broker. El borrowing rate de 0.0025 anual es bajo y puede aumentar durante períodos de estrés de mercado, incrementando costos de posiciones short.

El tamaño de muestra es limitado con solo 12 trades totales (9 + 3) ejecutados durante 12 años. Aunque las métricas son consistentes, un número mayor de trades proporcionaría mayor confianza estadística en las estimaciones. La distribución de retornos de trades individuales no sigue necesariamente una distribución normal, haciendo que métricas basadas en media y desviación estándar sean potencialmente engañosas.

El período de estudio incluye condiciones de mercado inusualmente favorables para farmacéuticas, particularmente la pandemia de COVID-19 (2020-2022) que generó volatilidad significativa en el sector. Períodos futuros podrían no replicar estas condiciones, afectando la frecuencia y magnitud de oportunidades de trading.

El overfitting a parámetros específicos ($\theta = 3.20$, $\varepsilon = 1.80$, $Q = 1e-5$, $R = 1e-2$) es un riesgo real. Aunque estos parámetros son razonables basándose en literatura y práctica común, fueron seleccionados con conocimiento implícito del período histórico. Una optimización exhaustiva mediante grid search podría haber resultado en overfitting a ruido histórico. La validación cruzada en series temporales es desafiante debido a dependencias temporales.

El modelo asume normalidad en los errores del Filtro de Kalman ($\omega_t \sim N(0, Q)$ y $v_t \sim N(0, R)$), pero retornos financieros típicamente exhiben colas pesadas (fat tails) y asimetría. Eventos extremos (black swans) ocurren con mayor frecuencia que lo predicho por distribuciones normales, exponiendo la estrategia a riesgos no capturados por el modelo.

Ventajas de los Filtros de Kalman

Los Filtros de Kalman demostraron ventajas claras sobre enfoques estáticos tradicionales en nuestro estudio. La adaptación automática a cambios de régimen fue evidente durante el período de test, donde eventos como la pandemia de COVID-19 alteraron fundamentalmente las dinámicas del sector farmacéutico. Un modelo estático con parámetros fijos de 2019 habría generado señales erróneas durante

2020-2022, pero el Filtro de Kalman ajustó $\beta_1(t)$ dinámicamente, manteniendo el spread centrado.

La cuantificación de incertidumbre mediante la matriz de covarianza $P(t)$ proporciona información valiosa sobre la confiabilidad de las estimaciones actuales. Durante períodos de alta incertidumbre, el sistema podría implementar controles de riesgo más estrictos, como reducir tamaños de posición. Durante períodos de baja incertidumbre, el sistema puede operar con mayor confianza.

El procesamiento online de datos es crucial para aplicaciones de trading real. El Filtro de Kalman procesa cada nueva observación secuencialmente sin necesidad de re-entrenar sobre todo el histórico. Esto reduce significativamente la carga computacional comparada con métodos batch como rolling regressions, permitiendo escalabilidad a portfolios de múltiples pares.

La robustez a outliers es una propiedad emergente del Kalman Gain adaptativo. Cuando una observación es altamente inconsistente con la predicción (innovación grande), el filtro reduce automáticamente su peso (K_t pequeño) si R es apropiadamente calibrado. Esto previene que observaciones anómalas corrompan las estimaciones de parámetros.

Aplicabilidad a Otros Pares

La metodología desarrollada es generalizable a otros pares de activos que cumplan las condiciones de cointegración. Sectores con características similares al farmacéutico son candidatos naturales. En el sector de utilities (empresas de electricidad, gas, agua), las empresas enfrentan regulaciones comunes, demanda relativamente inelástica, y estructuras de costos similares. Ejemplos incluirían Duke Energy vs Southern Company, o Dominion Energy vs Exelon.

El sector de telecomunicaciones presenta dinámicas similares con empresas respondiendo a regulaciones FCC, inversiones en infraestructura 5G, y competencia por market share. Pares como Verizon vs AT&T, o T-Mobile vs Sprint (pre-fusión) podrían exhibir cointegración explotable.

En commodities relacionados, pares como oro vs plata, petróleo WTI vs Brent, o gas natural vs petróleo crudo presentan relaciones de equilibrio de largo plazo basadas en fundamentos físicos y económicos. Estos pares son populares en trading de futuros y nuestro enfoque Kalman sería directamente aplicable.

El sector financiero ofrece numerosos candidatos. Bancos regionales comparables como Wells Fargo vs Bank of America, o instituciones de inversión como Goldman Sachs vs Morgan Stanley, comparten exposición a tasas de interés, regulaciones bancarias, y ciclos económicos. REITs (Real Estate Investment Trusts) especializados en el mismo segmento (retail, oficinas, multifamiliar) también podrían exhibir cointegración.

Sin embargo, es crucial realizar tests rigurosos de cointegración para cada par candidato. La intuición económica es necesaria pero no suficiente; la verificación estadística mediante ADF y Johansen en datos de entrenamiento es imperativa antes de desplegar capital real.

Extensiones y Mejoras Futuras

Varias extensiones podrían mejorar la estrategia presentada. La optimización de hiperparámetros mediante técnicas más sofisticadas que grid search simple podría incrementar performance. El algoritmo Expectation-Maximization (EM) estima Q y R directamente de los datos maximizando la verosimilitud del modelo. Bayesian optimization podría explorar eficientemente el espacio de parámetros (θ , ϵ , Q, R) minimizando el número de evaluaciones costosas.

La incorporación de múltiples pares en un portfolio diversificado reduciría el risk idiosincrático de operar un solo par. Un portfolio de 10-20 pares farmacéuticos, de utilities, de telecomunicaciones, etc., permitiría capturar oportunidades más frecuentemente mientras diversifica riesgo específico. La correlación entre pares debe gestionarse cuidadosamente para evitar concentración de riesgo.

El uso de vectores de Johansen para construir portfolios de $N > 2$ activos extendería el concepto de pairs trading a triplets o portfolios más grandes. El test de Johansen puede identificar múltiples vectores de cointegración en sistemas de N activos, cada uno representando una relación de equilibrio independiente explotable.

La implementación de controles de riesgo dinámicos basados en la incertidumbre del Filtro de Kalman mejoraría la gestión de capital. Cuando $P(t)$ es grande (alta incertidumbre), el sistema podría reducir tamaños de posición o incrementar umbrales (θ más grande). Cuando $P(t)$ es pequeña, el sistema podría operar más agresivamente.

La incorporación de machine learning para predicción de régimen podría complementar el enfoque Kalman. Modelos de clasificación podrían identificar períodos donde la cointegración es fuerte vs débil, activando o desactivando el

trading automáticamente. Features como volatilidad implícita, spreads de crédito, y sentimiento de noticias podrían mejorar estas predicciones.

El desarrollo de stop-loss dinámicos basados en la trayectoria de $P(t)$ y el z-score protegería contra breakdowns de cointegración. Si el spread excede un umbral extremo (ej. $|z| > 5$) y $P(t)$ está aumentando (creciente incertidumbre), el sistema podría cerrar posiciones preventivamente antes de pérdidas catastróficas.

La exploración de frecuencias mayores (intraday, horaria) adaptaría la metodología a horizontes de trading más cortos. Los Filtros de Kalman son particularmente adecuados para alta frecuencia debido a su naturaleza online. Sin embargo, costos de transacción y slippage se vuelven más significativos en frecuencias altas, requiriendo análisis cuidadoso de viabilidad económica.

Conclusiones

Este estudio demostró exitosamente la superioridad de Filtros de Kalman sobre métodos estáticos tradicionales en estrategias de Pairs Trading basadas en cointegración. La aplicación al par JNJ-PFE generó resultados excepcionales en términos de retorno ajustado por riesgo, validados rigurosamente mediante división temporal 60/20/20 y evaluación out-of-sample genuina.

La metodología desarrollada integra teoría econométrica clásica (Engle-Granger, Johansen) con técnicas modernas de estimación recursiva (Filtros de Kalman), creando un sistema que combina fundamentos teóricos sólidos con adaptabilidad práctica. El pipeline completo desde adquisición de datos hasta generación de señales y backtesting fue implementado con atención meticulosa a prevención de lookahead bias y realismo de ejecución.

Los resultados empíricos validan las hipótesis centrales del trabajo. El retorno anualizado de 37.92% en entrenamiento y 46.07% en test, ambos con Sharpe Ratio de 0.97, demuestran que la estrategia genera retornos substanciales ajustados por riesgo. El control de drawdown excepcional (máximo 1.03% en train, 0.49% en test) evidencia gestión de riesgo efectiva. Los Profit Factors extraordinarios (7846.51 en train, 234.08 en test) indican asimetría favorable entre ganancias y pérdidas.

La robustez out-of-sample es quizás el logro más significativo. La consistencia de métricas entre train y test, particularmente el Sharpe Ratio idéntico de 0.97, sugiere que los resultados no son producto de overfitting sino de captura genuina de una relación económica fundamental. La mejora de algunas métricas en test (Sortino, Calmar) sobre train es inusual y altamente favorable.

La baja frecuencia de trading (1 trade por año aproximadamente) hace que la estrategia sea accesible para traders individuales o instituciones pequeñas sin infraestructura de alta frecuencia. Los costos de transacción son manejables y la estrategia no requiere monitoreo intradía continuo, siendo compatible con operación part-time.

La transparencia e interpretabilidad del enfoque es una ventaja práctica significativa. A diferencia de modelos de machine learning tipo "caja negra", los Filtros de Kalman producen parámetros interpretables ($\beta_0(t)$, $\beta_1(t)$) cuya evolución puede monitorearse y entenderse. Esta interpretabilidad facilita debugging, mejora de confianza, y cumplimiento regulatorio en contextos institucionales.

Limitaciones Reconocidas

Es imperativo reconocer las limitaciones del estudio para contextualizar apropiadamente los resultados. El período de análisis (2010-2025) incluye un evento extremo (COVID-19) que generó volatilidad inusual en farmacéuticas. Períodos futuros sin eventos similares podrían mostrar menor frecuencia de oportunidades o magnitudes de reversión menores.

El número de trades es limitado (12 totales), proporcionando menos certeza estadística que estrategias de mayor frecuencia con miles de trades. Los intervalos de confianza alrededor de nuestras estimaciones de Sharpe Ratio, Sortino, etc., son amplios debido al tamaño de muestra pequeño.

Los supuestos del Filtro de Kalman (errores normales, modelo lineal) son aproximaciones que pueden no capturar completamente las dinámicas no-lineales y fat-tailed de mercados financieros reales. Extensiones como Extended Kalman Filter o Unscented Kalman Filter podrían abordar no-linealidades, pero agregan complejidad.

El backtest asume ejecución en precios de cierre sin considerar slippage intraday o impacto de mercado. En implementación real, especialmente con tamaños de posición grandes, estos factores erosionarían los retornos observados. Un modelo de costos de transacción más sofisticado sería necesario para deployment institucional.

Reflexión Final

Este trabajo demuestra que la intersección de teoría econométrica clásica, técnicas de ingeniería de control modernas, y implementación computacional cuidadosa puede generar estrategias de trading cuantitativo robustas y rentables. Los Filtros de Kalman, desarrollados originalmente para navegación espacial en los 1960s,

encuentran aplicación natural y efectiva en el problema de estimar relaciones dinámicas entre activos financieros.

La cointegración, concepto fundamental de econometría que valió a Engle y Granger el Premio Nobel en Economía, proporciona la fundamentación teórica que distingue pairs trading de gambling especulativo. La combinación de cointegración estática (identificando pares viables) con estimación dinámica Kalman (adaptándose a cambios graduales) representa un marco coherente y poderoso.

Los resultados obtenidos son alentadores pero deben interpretarse con humildad apropiada. Los mercados financieros son adversarios adaptativos; estrategias exitosas eventualmente se degradan debido a crowding, cambios estructurales, o simple reversión a la media de alfa. Vigilancia continua, adaptación, y gestión de riesgo disciplinada son imperativos para longevidad de cualquier estrategia cuantitativa.

Para estudiantes y practitioners de ingeniería financiera, este estudio ilustra la importancia de rigor metodológico. La división temporal estricta, la prevención meticulosa de lookahead bias, y la validación out-of-sample genuina son standards que toda investigación cuantitativa debe aspirar a cumplir. Resultados impresionantes in-sample son fáciles de obtener mediante overfitting; resultados que persisten out-of-sample representan conocimiento genuino explotable.

Finalmente, este trabajo subraya que trading cuantitativo exitoso requiere competencia multidisciplinaria. Conocimiento estadístico (cointegración, tests ADF), habilidades de programación (Python, manejo de datos, backtesting), entendimiento de mercados financieros (estructura de trading, costos, liquidez), y dominio de técnicas de optimización y control (Filtros de Kalman) deben converger coherentemente. La era moderna de quant trading favorece a equipos y individuos que integran estas disciplinas efectivamente.

Referencias

- Engle, R. F., & Granger, C. W. J. (1987).** "Co-integration and Error Correction: Representation, Estimation, and Testing." *Econometrica*, 55(2), 251-276.
- Johansen, S. (1991).** "Estimation and Hypothesis Testing of Cointegration Vectors in Gaussian Vector Autoregressive Models." *Econometrica*, 59(6), 1551-1580.
- Kalman, R. E. (1960).** "A New Approach to Linear Filtering and Prediction Problems." *Journal of Basic Engineering*, 82(1), 35-45.
- Gatev, E., Goetzmann, W. N., & Rouwenhorst, K. G. (2006).** "Pairs Trading: Performance of a Relative-Value Arbitrage Rule." *Review of Financial Studies*, 19(3), 797-827.
- Vidyamurthy, G. (2004).** *Pairs Trading: Quantitative Methods and Analysis*. Wiley Finance.
- Libro práctico que cubre implementación de estrategias de pares desde perspectiva de practitioner, incluyendo código y ejemplos.