



UNIVERSITAT DE
BARCELONA

Facultat de Matemàtiques
i Informàtica

**Doble grado en Matemáticas y
Administración y Dirección de Empresas**

Trabajo Final de Grado

**MERCADO EFICIENTE, MAPAS
AUTOORGANIZADOS Y SUS
APLICACIONES EN LAS FINANZAS**

Autor: Genis Guglielminotti de Inés

**Director: Dr. Jose Manuel Corcuera Valverde
Dr. Salvador Torra i Porras**

**Realitzat a: Departament de Matemàtiques i Informàtica
Departament d'Econometria, Estadística i
Economia Aplicada**

Barcelona, 24 de enero de 2022

Abstract

The aim of this paper is to study the efficient market fundamentals, to analyse in detail the technique of self-organising maps and to see how both tools complement each other to create efficient portfolios.

Part I introduces the efficient market and Markowitz theory about portfolios optimization. We will also the market indices, their behaviour, construction and usefulness in the investment field. We will think about possible ways to optimise them and we will learn about the SP&500, the most important index at present.

Part II studies how Self-Organising Maps work and how they are able to reduce a large volume of data to smaller dimensions, taking an in-depth look at the algorithmic part behind this technique.

Part III seeks to create efficient or optimal portfolios using the self-organised maps, working in an environment constrained by the assets that compose the SP&500 and compare them.

Resumen

El objetivo de este trabajo es estudiar las bases del mercado eficiente, analizar con detalle la técnica de los mapas autoorganizados y ver como ambas herramientas se complementan para crear carteras eficientes.

La parte I presenta el mercado eficiente y la teoría de Markowitz acerca de la optimización de carteras. También descubriremos qué son los índices bursátiles, su comportamiento, construcción y utilidad que tienen en el campo de la inversión. Pensaremos posibles maneras de optimizarlos y conoceremos el SP&500, el índice más importante en la actualidad.

La parte II estudia el funcionamiento de los mapas autoorganizados y como éstos son capaces de reducir un gran volumen de datos a dimensiones menores, analizando profundamente la parte algorítmica que hay detrás de esta técnica.

La parte III busca crear carteras eficientes u óptimas utilizando los mapas autoorganizados, trabajando en un entorno acotado por los activos que componen el SP&500 y compararlos.

Agradecimientos

Agradecer al Dr. Jose Manuel Corcuera y al Dr. Salvador Torra la supervisión y los consejos que me han dado para llevar a cabo esta memoria.

Agradecer también a mi amigo y matemático Victor Valldosera por ayudarme desinteresadamente en todo lo que he necesitado para la realización de este trabajo.

Finalmente, mi más sincero agradecimiento a todas aquellas personas, después compañeros y ahora ya amigos, que he tenido el placer de conocer a lo largo de estos años. Sin vosotros nada de esto habría sido posible.

Índice

I	Estudio del mercado eficiente	1
1.	Mercado eficiente	2
1.1.	Marco teórico	2
1.2.	Carteras de inversión	3
1.3.	Modelo de Markowitz	4
1.3.1.	Frontera eficiente y portafolio de mínima varianza	6
1.3.2.	Teorema de dos fondos	8
1.3.3.	Teorema de un fondo	8
2.	Índices bursátiles	10
2.1.	Introducción	10
2.2.	Tipos de índices	10
2.2.1.	Índices por capitalización bursátil	11
2.2.2.	Índices por precio bursátil	12
2.2.3.	Índices por igualdad de pesos	13
2.2.4.	Índices alternativos	13
2.3.	Construcción	14
2.3.1.	Sesgo muestral	14
2.3.2.	Sesgo en la construcción	14
2.3.3.	Error de seguimiento	14
2.3.4.	Sesgo de eficiencia	15
2.3.5.	Sesgo activo	15
2.4.	Propiedades	15
2.5.	Indicador de Herfindhal	15
2.6.	Ratios financieras	16
2.7.	Exchange Trade Funds (ETF's)	17
3.	S&P 500	18
3.1.	Requisitos	19
3.2.	Propiedades	19
3.3.	Sectores	20
II	Estudio de la técnica de los mapas autoorganizados	21

4. Mapas autoorganizados	22
4.1. Introducción	22
4.2. Modelos de aprendizaje	22
4.3. Definición	23
4.4. Estructura	23
4.5. Variables	24
4.6. Algoritmo de aprendizaje	25
4.6.1. Inicialización	25
4.6.2. Elección	26
4.6.3. Competición	26
4.6.4. Actualización	26
4.7. Propiedades	27
III Construcción de una cartera óptima	28
5. Introducción	29
6. Construcción de las carteras	29
6.1. Base de datos	29
6.2. Creación del mapa autoorganizado	30
6.3. Selección de activos	31
6.4. Constitución de las carteras	33
6.5. Representación de la frontera eficiente	34
7. Valoración de las carteras	35
8. Conclusiones	36

Parte I

Estudio del mercado eficiente

La parte I explica los conceptos fundamentales del mercado eficiente, explicando primeramente qué es el mercado de valores y presentando conceptos financieros que van a ser importantes para el desarrollo del problema de Markowitz. En segunda instancia profundizaremos en el tratamiento de los índices bursátiles: qué son, qué representan, cómo se construyen y porqué son útiles.

1. Mercado eficiente

1.1. Marco teórico

Definición 1.1. *El mercado de valores es el lugar donde se intercambian diferentes tipos de activos o productos financieros cuyo precio cambia en el tiempo y viene determinado por la oferta y la demanda.*

Definición 1.2. *Llamamos **inversor** a la persona, ya bien sea física o jurídica, que acude al mercado para realizar una acción de compra o venta de valores con el objetivo de generar unos rendimientos a su patrimonio en el corto, medio o largo plazo.*

Definición 1.3. *Llamamos **broker** al organismo intermediario entre el mercado y el inversor cuya función principal es ejecutar cualquier operación de compra o venta que desee hacer el inversor, cobrando una comisión por ello.*

Definición 1.4. *El **mercado eficiente** es aquel cuyos precios reflejan perfectamente toda la información del mercado: cualquier persona tiene acceso gratuito a la toda la información relevante del mercado, no existen acciones infravaloradas o sobrevaloradas ya que el precio se actualiza con la información y hay muchos participantes en el mercado.*

Las tres propiedades que definen la calidad de un activo financiero son:

Definición 1.5. *La **rentabilidad** es la capacidad que tiene un activo de generar rendimientos futuros.*

Definición 1.6. *El **riesgo** es la incertidumbre asociada a un activo para predecir su rendimiento futuro.*

Definición 1.7. *La **liquidez** es la cualidad que tiene un activo para poder transformarse fácilmente en dinero efectivo.*

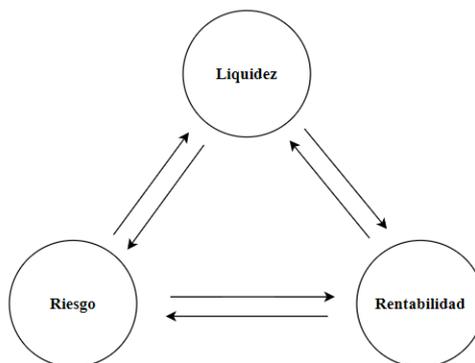


Figura 1.1.1 (Elaboración Propia)

El dilema del inversor es elegir las dos variables que quiere priorizar y a cual de estas quiere renunciar a la hora de realizar una inversión para escoger un tipo de activo u otro. Los activos inmobiliarios son un ejemplo de activos con buena rentabilidad y poco riesgo, pero tienen poca liquidez. En cambio, los activos de renta variable tienen mucha rentabilidad y muy líquidos pero mucho riesgo.

1.2. Carteras de inversión

Las carteras de inversión son conjuntos de activos a los que se les da un peso determinado y que generan un rendimiento para el inversor, que es el responsable de determinar cuáles serán aquellos activos que compondrán la cartera. Aunque el tipo de activos que podemos incluir en una cartera sea muy variado, en este trabajo solo hablaremos de renta variable o acciones. Para poder estudiar bien lo que es una cartera de inversión, es necesario presentar los siguientes conceptos financieros:

Definición 1.8. Sea n el número de acciones que tiene una compañía y p el precio al que cotiza cada acción, el valor de mercado o **capitalización bursátil** de una compañía es el dinero que pagaría un inversor si quisiera comprar toda la compañía a través de acciones en el mercado tal que

$$V_i = n \cdot p.$$

Es importante tener en cuenta que este valor es variante en el tiempo ya que la cotización de las acciones de una empresa en el mercado cambian continuamente.

Definición 1.9. Sea p_0 el precio del activo en el inicio del periodo y p_f el precio del activo al final del periodo, la **rentabilidad** de un activo mide el rendimiento tal que

$$r_i = \left[\frac{p_f - p_0}{p_0} \right] \cdot 100.$$

Notar que la rentabilidad de un activo hace referencia a un periodo determinado sin tener en cuenta la duración del mismo, pues es posible calcular una rentabilidad mensual, trimestral, anual, bienal, etc.

Definición 1.10. Sea p_0 el precio del activo en el inicio del periodo y p_m el precio del activo al cabo de m años, la **tasa de crecimiento anual compuesto** o **tasa anualizada de retorno** (en inglés, *CAGR* o *compound annual growth rate*) determina el rendimiento anual de un activo tal que

$$CAGR_i = \left[\left(\frac{p_m}{p_0} \right)^{\frac{1}{m}} - 1 \right] \cdot 100.$$

Esta variable es imprescindible para valorar el rendimiento de una cartera ya que esta tasa de retorno anual indica si una inversión está generando un buen rendimiento o no entre dos momentos diferentes en el tiempo, reinvertiendo los beneficios al final de cada año.

Definición 1.11. Sea n el número de activos de la cartera y w_i el peso de cada uno de ellos, el **rendimiento anualizado del portafolio** es

$$r = \sum_{i=1}^n CAGR_i \cdot w_i.$$

Definición 1.12. Sea m el número de periodos, r_j la rentabilidad en el periodo j y \bar{r} la rentabilidad media de estos periodos, la **desviación típica** de un activo mide el riesgo o volatilidad tal que

$$\sigma_i = \sqrt{\frac{\sum_{j=1}^m (r_j - \bar{r})^2}{m - 1}}.$$

Desde un punto de vista matemático, la desviación típica mide el grado de dispersión de los puntos de una muestra respecto a su media. Es por esto que si un activo tiene rendimientos muy diferentes de un periodo a otro, el valor de esta variable será alto reflejando el riesgo o volatilidad.

Definición 1.13. Sea m el número de periodos, A y B dos activos diferentes de una cartera, r_j la rentabilidad en el periodo j y \bar{r} la rentabilidad media de estos periodos, la **correlación** entre dos activos A y B mide el grado de similitud que existe entre ambos tal que

$$\rho_{AB} = \frac{\left[\sum_{j=1}^m (r_j^A - \bar{r}^A) \cdot (r_j^B - \bar{r}^B) \right]}{\sqrt{\sum_{j=1}^m (r_j^A - \bar{r}^A)^2 \cdot \sum_{j=1}^m (r_j^B - \bar{r}^B)^2}}.$$

Definición 1.14. Sea m el número de periodos, A y B dos activos diferentes de una cartera, r_j la rentabilidad en el periodo j y \bar{r} la rentabilidad media de estos periodos, la **covarianza** entre dos activos A y B mide el grado de variación conjunta respecto a sus medias tal que

$$\sigma_{AB} = \frac{\sum_{j=1}^m (r_j^A - \bar{r}^A) \cdot (r_j^B - \bar{r}^B)}{m - 1} = \rho_{AB} \cdot \sigma_A \cdot \sigma_B.$$

Definición 1.15. Sea n el número de activos de la cartera y w_i el peso de cada uno de ellos, la **desviación típica del portafolio** mide el riesgo total de la cartera tal que

$$\sigma = \sqrt{\sum_{i=1}^n w_i^2 \cdot \sigma_i^2 + 2 \sum_{\substack{i=1 \\ j>i}}^n \sigma_{ij} \cdot w_i \cdot w_j}.$$

Otra forma de calcular el riesgo del portafolio y que es una consecuencia directa del resultado anterior es a través de la varianza:

$$Var(p) = \sigma^2 = \sum_{i=1}^n w_i^2 \cdot \sigma_i^2 + 2 \sum_{\substack{i=1 \\ j>i}}^n \sigma_{ij} \cdot w_i \cdot w_j.$$

1.3. Modelo de Markowitz

Harry Markowitz, en el año 1952, propuso un modelo para poder construir un portafolio eficiente a partir del riesgo y rentabilidad. Como el rendimiento futuro de un activo i es incierto, el modelo supone $\mathbb{E}(r_i) := \bar{r}_i$ (ver Karl Sigman, 2005 [1]). De este modo, definimos la tasa esperada de retorno de una cartera como la media de las rentabilidades históricas de los activos que la componen:

$$\mathbb{E}(r) := \bar{r} = \sum_{i=1}^n \bar{r}_i \cdot w_i.$$

La optimización del par (\bar{r}, σ) dado un nivel de rentabilidad \bar{r} determinado, pasa por minimizar

$$\sigma^2 = \sum_{i=1}^n w_i^2 \cdot \sigma_i^2 + 2 \sum_{\substack{i=1 \\ j>i}}^n \sigma_{ij} \cdot w_i \cdot w_j,$$

dadas las siguientes dos restricciones:

$$\sum_{i=1}^n \bar{r}_i \cdot w_i = \bar{r},$$

$$\sum_{i=1}^n w_i = 1.$$

El método que propone Markowitz para resolver este problema es a través de los multiplicadores de Lagrange. Usando los multiplicadores λ y μ (uno por cada restricción que tenemos), obtenemos

$$L(w_1, \dots, w_n, \lambda, \mu) = \frac{1}{2} \sum_{i,j=1}^n \sigma_{ij} \cdot w_i \cdot w_j - \lambda \left(\sum_{i=1}^n \bar{r}_i \cdot w_i - \bar{r} \right) - \mu \left(\sum_{i=1}^n w_i - 1 \right),$$

que es la función que queremos optimizar. Así pues, fijando $\frac{\partial L}{\partial w_i} = \frac{\partial L}{\partial \lambda} = \frac{\partial L}{\partial \mu} = 0$ para los n activos, obtenemos el sistema de $n + 2$ ecuaciones

$$\begin{cases} \sum_{j=1}^n \sigma_{ij} \cdot w_j - \lambda \cdot \bar{r}_i - \mu = 0, & \forall i = 1, \dots, n, \\ \sum_{i=1}^n \bar{r}_i \cdot w_i = \bar{r}, \\ \sum_{i=1}^n w_i = 1, \end{cases} \quad (1.3.1)$$

que es lineal respecto $(w_1, \dots, w_n, \lambda, \mu)$. Ahora el problema se reduce en resolver este sistema de ecuaciones y encontrar una solución (w_1, \dots, w_n) utilizando distintos métodos numéricos. De este modo, el par (σ, \bar{r}) queda totalmente determinado por el riesgo predefinido y

$$\sigma = \sqrt{\sum_{i,j=1}^n \sigma_{ij} \cdot w_i \cdot w_j}.$$

De este modo es claro ver como Markowitz propone un modelo que determina el riesgo de un portafolio no solo a través de la varianza de cada activo sino también a través de la covarianza entre ellos, obteniendo un menor riesgo cuando la combinación de los activos consigue un buen nivel de diversificación.

1.3.1. Frontera eficiente y portafolio de mínima varianza

Definición 1.16. El *conjunto factible* es una región infinita donde se ubican todas las carteras posibles, es decir, aquellas $w = (w_1, \dots, w_n)$ tales que $\sum_{i=1}^n w_i = 1$.

Definición 1.17. Las *carteras eficientes* son aquellas $w = (w_1, \dots, w_n)$ que satisfacen el problema (1.3.1) dado un nivel de riesgo $\bar{r} \in \mathbb{R}$ y que, por lo tanto, están óptimamente diversificadas.

Existen infinitud de portafolios eficientes, uno por cada nivel de riesgo. La posibilidad de poder representar todas estas carteras en una gráfica da lugar a la frontera eficiente.

Definición 1.18. Llamamos *frontera eficiente* o *conjunto de mínima varianza* a todas aquellas carteras o pares (σ, \bar{r}) tales que, dado un nivel de riesgo determinado, obtienen la mínima varianza.

Definición 1.19. El *portafolio de mínima varianza* (σ^*, \bar{r}^*) es aquel que tiene la menor varianza de entre todos los que se encuentran en la frontera eficiente.

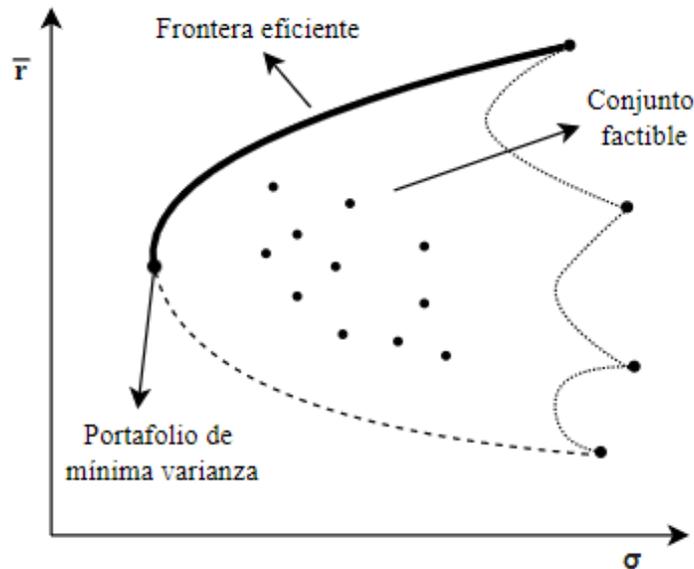


Figura 1.3.1.1 (Elaboración Propia)

La frontera eficiente acota superiormente el conjunto factible, ya que por encima de esta no existen carteras posibles. El número de activos que componen la cartera, siempre y cuando sean dos o más, no afecta a la forma de la frontera eficiente que es esta especie de parábola invertida, pero sí al conjunto factible, que tiene tantos picos como activos tiene la cartera (la figura representa un portafolio compuesto por cuatro activos).

De toda la curva representada, en la práctica no consideramos la parte inferior cuando hablamos de la frontera eficiente. Eso es debido a la existencia de portafolios con misma varianza y diferente rentabilidad, pues el inversor racional siempre escogerá aquella cuya rentabilidad sea mayor, que es la que se sitúa en la parte superior de la parábola.

Corolario 1.20. *Dado un número finito de activos, el portafolio de mínima varianza existe y es único.*

Demostración. Siguiendo con el modelo de Markowitz, queremos minimizar

$$\sigma^2 = \sum_{i=1}^n w_i^2 \cdot \sigma_i^2 + 2 \sum_{\substack{i=1 \\ j>i}}^n \sigma_{ij} \cdot w_i \cdot w_j$$

que, como $\sigma_{ii} = \sigma_i^2$, equivale a

$$\frac{1}{2} \sum_{i,j=1}^n \sigma_{ij} \cdot w_i \cdot w_j.$$

En este caso, únicamente tenemos la restricción

$$\sum_{i=1}^n w_i = 1,$$

ya que aquí no suponemos una rentabilidad determinada inicial. De modo que al aplicar el método de Lagrange solo aparece el multiplicador μ en la función que queremos maximizar

$$L(w_1, \dots, w_n, \mu) = \frac{1}{2} \sum_{i,j=1}^n \sigma_{ij} \cdot w_i \cdot w_j - \mu \left(\sum_{i=1}^n w_i - 1 \right).$$

Fijando $\frac{\partial L}{\partial w_i} = \frac{\partial L}{\partial \mu} = 0$ para todos los n activos, obtenemos el sistema de $n + 1$ ecuaciones

$$\begin{cases} \sum_{j=1}^n \sigma_{ij} \cdot w_j - \mu = 0, & \forall i = 1, \dots, n, \\ \sum_{i=1}^n w_i = 1, \end{cases}$$

que es lineal respecto (w_1, \dots, w_n, μ) y que equivale a resolver la ecuación matricial (ver Eric Zivot, 2013 [2]) siguiente:

$$\underbrace{\begin{bmatrix} 2\sigma_1^2 & \dots & 2\sigma_{1n} & 1 \\ \vdots & \ddots & \vdots & \vdots \\ 2\sigma_{n1} & \dots & 2\sigma_n^2 & 1 \\ 1 & \dots & 1 & 0 \end{bmatrix}}_A \begin{bmatrix} w_1 \\ \vdots \\ w_n \\ \mu \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix}. \quad (1.1)$$

Al ser la matriz A invertible, la solución

$$\begin{bmatrix} w_1 \\ \vdots \\ w_n \\ \mu \end{bmatrix} = A^{-1} \cdot \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix} \quad (1.2)$$

es única.

□

1.3.2. Teorema de dos fondos

Si bien es cierto que resolviendo el problema de Markowitz para infinitos niveles de riesgo es posible representar la frontera eficiente, existe un teorema que establece el número mínimo de soluciones del problema de Markowitz que es necesario para poder construir la frontera eficiente en su totalidad.

Teorema 1.21 (teorema de dos fondos). *Dadas dos soluciones diferentes del problema de Markowitz se pueden generar infinitas nuevas soluciones, encontrándose todas estas en la frontera eficiente.*

Demostración. Sean

$$\begin{cases} w^1 = (w_1^1, \dots, w_n^1, \lambda^1, \mu^1) & \text{dado } \bar{r}^1, \\ w^2 = (w_1^2, \dots, w_n^2, \lambda^2, \mu^2) & \text{dado } \bar{r}^2 \end{cases}$$

soluciones diferentes del problema (1.3.1). Para cualquier $0 < \alpha < 1$ tenemos

$$\begin{cases} \alpha w^1 = (\alpha w_1^1, \dots, \alpha w_n^1, \alpha \lambda^1, \alpha \mu^1), \\ (1 - \alpha) w^2 = ((1 - \alpha) w_1^2, \dots, (1 - \alpha) w_n^2, (1 - \alpha) \lambda^2, (1 - \alpha) \mu^2). \end{cases}$$

Entonces, la cartera $w^3 = \alpha w^1 + (1 - \alpha) w^2$ cumple las siguientes ecuaciones:

$$\begin{aligned} \sum_{i=1}^n \bar{r}_i \cdot w_i^3 &= \sum_{i=1}^n \bar{r}_i [\alpha w_i^1 + (1 - \alpha) w_i^2] = \alpha \sum_{i=1}^n \bar{r}_i \cdot w_i^1 + (1 - \alpha) \sum_{i=1}^n \bar{r}_i \cdot w_i^2 = \alpha \bar{r}^1 + (1 - \alpha) \bar{r}^2 = \bar{r}^3 \\ \sum_{i=1}^n w_i^3 &= \sum_{i=1}^n [\alpha w_i^1 + (1 - \alpha) w_i^2] = \alpha \sum_{i=1}^n w_i^1 + (1 - \alpha) \sum_{i=1}^n w_i^2 = \alpha + (1 - \alpha) = 1. \end{aligned}$$

Tomando $\lambda^3 = \alpha \lambda^1 + (1 - \alpha) \lambda^2$ y $\mu^3 = \alpha \mu^1 + (1 - \alpha) \mu^2$, obtenemos $w^3 = (w_1^3, \dots, w_n^3, \lambda^3, \mu^3)$ como solución del problema (1.3.1), cuyo nivel de riesgo asociado es $\bar{r}^3 = \alpha \bar{r}^1 + (1 - \alpha) \bar{r}^2$. \square

Por lo tanto, a partir de dos soluciones diferentes del problema de Markowitz es posible encontrar todas las otras soluciones que determinan, en su totalidad, la frontera eficiente. Si pensamos las dos soluciones distintas como oportunidades de inversión (como si de fondos de inversión se trataran), no es necesario ningún otro activo o fondo para poder llevar a cabo una inversión óptima para el inversor racional sino encontrar la combinación entre ellos que se ajuste a sus necesidades.

1.3.3. Teorema de un fondo

Si bien es cierto que hasta ahora solo hemos tratado activos con un supuesto nivel de riesgo, existen los activos libres de riesgo que también se pueden incorporar en un portafolio. Como indica su propio nombre, su varianza o riesgo es nulo y tiene un rendimiento r_f . De aquí surge la posibilidad de crear una nueva cartera a partir de un portafolio de activos con riesgo y el activo libre de riesgo que acabamos de presentar (ver Karl Sigman, 2005 [3]).

De ahora en adelante, notaremos el portafolio compuesto solamente por los activos de riesgo A_1, \dots, A_n como $w = (w_1, \dots, w_n)$, y el portafolio incluyendo el activo libre de riesgo A_0 como $v = (v_0, v_1, \dots, v_n) = (v_0, (1 - v_0)(w_1, \dots, w_n))$, que podemos tratar como dos fondos independientes, de modo que

$$\begin{cases} \bar{r}_w = \sum_{i=1}^n w_i \cdot \bar{r}_i, \\ \sigma_w = \sqrt{\sum_{i,j=1}^n \sigma_{ij} \cdot w_i \cdot w_j}, \end{cases}$$

y

$$\begin{cases} \bar{r}_v = v_0 \cdot r_f + (1 - v_0) \cdot \bar{r}_w, \\ \sigma_v = (1 - v_0) \cdot \sigma_w, \end{cases}$$

obteniendo pues

$$(\sigma_v, \bar{r}_v) = ((1 - v_0) \cdot \sigma_w, v_0 \cdot r_f + (1 - v_0) \cdot \bar{r}_w) = (1 - v_0)(\sigma_w, \bar{r}_w) + v_0(0, r_f).$$

Como $0 < v_0 < 1$, (σ_v, \bar{r}_v) se mueve en la línea

$$\bar{r}_v = \frac{\bar{r}_w - r_f}{\sigma_w} \sigma_v + r_f.$$

Definición 1.22. *La nueva frontera eficiente de la cartera $v = (v_0, v_1, \dots, v_n)$ o línea del mercado de capitales es el segmento que conecta el punto $(0, r_f)$ y $F = (\sigma_w, \bar{r}_w)$ que es tangente a la antigua frontera eficiente dibujada por la cartera $w = (w_1, \dots, w_n)$.*

Definición 1.23. *Llamamos cartera del mercado a aquella $v = (v_0, v_1, \dots, v_n)$ que combina activos libres de riesgo y activos con riesgo que se sitúa en $F = (\sigma_w, \bar{r}_w)$.*

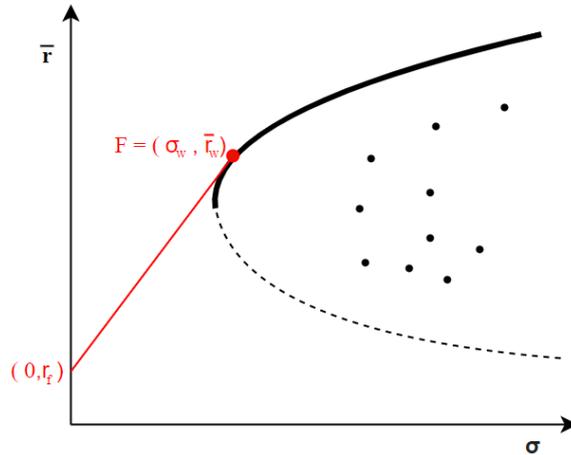


Figura 1.3.3.1 (Elaboración Propia)

Teorema 1.24 (teorema de un fondo). *Cualquier portafolio que se sitúa en la nueva frontera eficiente puede ser creado como combinación de la cartera $w = (w_1, \dots, w_n)$ determinada por $F = (\sigma_w, \bar{r}_w)$, que es única, y el activo libre de riesgo.*

2. Índices bursátiles

La teoría de índices bursátiles surge a finales del siglo XIX y explica el comportamiento de los mercados desde una perspectiva financiera con el objetivo de ayudar al inversor y proporcionarle la mayor cantidad de información posible. La aparición de nuevos índices en el mercado es evidente, siendo cada vez mayor los productos financieros que se basan en ellos como las opciones y los futuros.

2.1. Introducción

Definición 2.1. *Un índice bursátil es un indicador de referencia que trata de representar el comportamiento del mercado, o una parte de él, a través de las empresas que lo forman.*

El objetivo de un índice no es construir una cartera que sea rentable para el inversor (que lo puede ser, por ejemplo, en la inversión pasiva), sino representar el comportamiento del mercado de la manera más fiel posible. Es por esto que, a través de los movimientos de un índice bursátil, podemos extraer información del mercado que ese índice representa y que el inversor puede utilizar para valorar sus carteras de inversión, ofreciendo la posibilidad de poder comparar rendimientos y realizar cambios en la composición de sus carteras.

Para la correcta comprensión de las próximas secciones, es imprescindible presentar los siguientes conceptos:

Definición 2.2. *El valor de un índice es el cociente entre el valor individual de todas las compañías (que varía en función del método de ponderación escogido) y el divisor.*

Definición 2.3. *El divisor de un índice es una variable que se va ajustando continuamente para compensar posibles cambios en el valor de mercado de un índice, y así evitar que el índice sufra grandes variaciones en un corto periodo de tiempo.*

En la creación de un índice, se establece un valor inicial al divisor en función del valor total de las compañías que lo componen para que el valor del índice sea mucho más fácil de controlar. El valor del divisor cambiará cuando se produzcan algunas de las siguientes operaciones: cambios en la composición del índice, emisión de nuevas acciones, fusiones, venta parcial de la compañía, reparto extraordinario de dividendos, etc. El ajuste del divisor se realiza al cierre del día, garantizando así que el valor del índice será igual en la apertura de la jornada siguiente si los precios se mantienen. Por lo tanto, el valor del índice cambiará únicamente en función de las fluctuaciones del mercado y cambios en el precio de las acciones, caso en que no se produce una actualización del divisor.

2.2. Tipos de índices

En función de la ponderación de los activos, los índices pueden tener rendimientos o comportamientos totalmente distintos. La actualización de los divisores, que es clave para determinar como se construyen, también sufre modificaciones en función de un tipo de índice u otro. En este apartado trabajaremos distintos tipos de índices en función de la ponderación de sus activos y como estos consiguen mantener el valor en el proceso de actualización (ver [4] y [5]).

2.2.1. Índices por capitalización bursátil

Definición 2.4. Sea n el número de activos que componen el índice, p_i el precio que cotiza la acción y q_i el número de acciones negociables que tiene la compañía en el mercado, el **peso o ponderación** de un activo dentro del índice es

$$w_i = \frac{p_i \cdot q_i}{\sum_{j=1}^n p_j \cdot q_j}.$$

Este tipo de ponderación se llama capitalización bursátil del capital flotante ya que solo tiene en cuenta el número de acciones que están disponibles en el mercado, excluyendo aquellas que no son negociables.

Valor del índice

Definición 2.5. Sea n el número de activos que componen el índice, p_i el precio que cotiza la acción, q_i el número de acciones negociables que tiene la compañía en el mercado y w_i la ponderación de cada activo, el **valor de un índice por capitalización bursátil** es

$$Valor = \frac{\sum_{i=1}^n p_i \cdot q_i \cdot w_i}{Divisor}.$$

Ajuste del divisor

Supongamos un cambio en la composición de la cartera, y un activo que forma parte de la composición del índice en el momento t es sustituido por otro activo que forma parte de la composición del índice en el momento $t + 1$, entonces

$$Valor = \frac{\sum_{i=1}^n p_i^t \cdot q_i^t \cdot w_i^t}{Divisor_t}$$

y

$$Valor = \frac{\sum_{i=1}^n p_i^{t+1} \cdot q_i^{t+1} \cdot w_i^{t+1}}{Divisor_{t+1}}.$$

Igualando tenemos

$$\frac{\sum_{i=1}^n p_i^t \cdot q_i^t \cdot w_i^t}{Divisor_t} = \frac{\sum_{i=1}^n p_i^{t+1} \cdot q_i^{t+1} \cdot w_i^{t+1}}{Divisor_{t+1}}$$

y, aislando, obtenemos

$$Divisor_{t+1} = Divisor_t \cdot \frac{\sum_{i=1}^n p_i^{t+1} \cdot q_i^{t+1} \cdot w_i^{t+1}}{\sum_{i=1}^n p_i^t \cdot q_i^t \cdot w_i^t} = \frac{\sum_{i=1}^n p_i^{t+1} \cdot q_i^{t+1} \cdot w_i^{t+1}}{Valor}.$$

2.2.2. Índices por precio bursátil

Definición 2.6. Sea n el número de activos que componen el índice y p_i el precio que cotiza la acción, el **peso** o **ponderación** de un activo dentro del índice es

$$w_i = \frac{p_i}{\sum_{j=1}^n p_j}.$$

A diferencia de la ponderación por capitalización, aquí solo se tiene en cuenta el precio que cotiza cada acción ignorando pues el número de acciones.

Valor del índice

Definición 2.7. Sea n el número de activos que componen el índice, p_i el precio que cotiza la acción y w_i la ponderación de cada activo, el **valor de un índice por precio bursátil** es

$$Valor = \frac{\sum_{i=1}^n p_i \cdot w_i}{Divisor}.$$

Ajuste del divisor

Siguiendo el ejemplo anterior y suponiendo que hay un cambio en la composición del índice, donde una compañía en el momento t es cambiada por otra en el momento $t + 1$, el valor del divisor actualizado queda determinado por la ecuación

$$Divisor_{t+1} = Divisor_t \cdot \frac{\sum_{i=1}^n p_i^{t+1} \cdot w_i^{t+1}}{\sum_{i=1}^n p_i^t \cdot w_i^t} = \frac{\sum_{i=1}^n p_i^{t+1} \cdot w_i^{t+1}}{Valor}.$$

En este caso, el divisor se ajusta cuando existen operaciones que afecten al precio de la acción que no estén relacionadas con el flujo del mercado: reparto extraordinario de dividendos, división de acciones o ajuste de precios. En cambio, la emisión de nuevas acciones no provoca una actualización del divisor ya que solo afecta al número de acciones que aquí no tenemos en cuenta.

Este tipo de ponderación es muy útil para aquellos índices que están compuestos por compañías que pertenecen al mismo sector y que tratan de representar una industria en concreto, ya que da una idea aproximada de la evolución de los precios del mercado. De este modo, aumenta la representación de esas compañías de menor capitalización que no tienen un gran número de acciones en circulación pero sí una gran evolución en el precio de cotización.

2.2.3. Índices por igualdad de pesos

Definición 2.8. Sea n el número de activos que componen el índice, el **peso** o **ponderación** de un activo dentro del índice es

$$w_i = \frac{1}{n}.$$

Valor del índice

Definición 2.9. Sea n el número de activos que componen el índice, p_i el precio que cotiza la acción y q_i el número de acciones negociables que tiene la compañía en el mercado, el **valor de un índice por igualdad de pesos** es

$$Valor = \frac{\sum_{i=1}^n p_i \cdot q_i}{n \cdot Divisor}.$$

Ajuste del divisor

Esta actualización es más compleja que en el caso anterior ya que también de debe ajustar el divisor en el caso de cambio del precio de las acciones para garantizar que el peso de cada activo es el que corresponde. En caso contrario, el valor del índice cambiaría en proporciones diferentes al peso de los activos. La actualización que se lleva a cabo es periódica en función de la política de cada índice.

La ecuación que define el nuevo valor del divisor después del rebalanceo por ajuste de precios es:

$$Divisor_{t+1} = \frac{Z \cdot (\sum_{i=1}^n p_i^{t+1} \cdot q_i^{t+1})}{n \cdot Valor},$$

donde Z es una constante específica del índice, que recalcula el número de acciones de cada compañía.

2.2.4. Índices alternativos

Es muy frecuente que, al calcular las ponderaciones por capitalización bursátil del capital flotante, algunas compañías representen un peso muy elevado dentro de la cartera y que puede ser contraproducente para los intereses del propio índice. En estos casos, existe una alternativa que consiste en establecer un límite máximo de ponderación para cada activo. De esta manera, si el activo con mayor capitalización bursátil supera este límite, se le otorga el peso determinado y las otras compañías se reparten el peso restante. En caso de que vuelva a suceder lo mismo, se reitera el proceso hasta que sea necesario.

2.3. Construcción

En el proceso de construcción de un índice se producen diferentes situaciones que permiten explicar, o entender, el comportamiento que tienen los índices y cuantificar el grado de error en el que incurrimos (ver Jordi Andreu Corbatón, 2009 [6]).

2.3.1. Sesgo muestral

Definición 2.10. *Llamamos **sesgo muestral** a la diferencia entre la rentabilidad de un índice y la rentabilidad del mercado que el índice representa.*

Viene dado por la imposibilidad de seleccionar todos los activos del mercado para un índice. La selección de una muestra en concreto puede tener un impacto negativo en la significación estadística del modelo, provocando pues la aparición de estimadores sesgados. Se puede corregir parte de este sesgo asignando cierta ponderación a los subgrupos no representados, aunque esta práctica podría incrementar el sesgo de otro tipo. Tener en cuenta que el capital flotante de las compañías ayuda a minimizar esta variable, excluyendo pues aquellas acciones que no afectan al comportamiento del mercado. La subjetividad en la elección de activos que no tienen nada que ver con la representación también afectan al sesgo muestral, como las políticas económicas o beneficios fiscales, ya que pueden influir a las decisiones tomadas por el inversor.

2.3.2. Sesgo en la construcción

Definición 2.11. *Llamamos **sesgo de construcción** a la diferencia entre la rentabilidad del índice (determinado por la ponderación de los activos) y la rentabilidad total del mercado.*

La rentabilidad total del mercado se determina de la siguiente manera: se compara el valor de mercado entre dos momentos del tiempo y se añaden los beneficios que comporta invertir en el portafolio y que no afectan al precio (derechos de oferta o dividendos.).

Teóricamente, el índice debería estar ponderado por capitalización bursátil para ser una buena aproximación del mercado, y todo lo que sea alejarse de este tipo de ponderación incrementa el sesgo por construcción. En la práctica, se aplican muchos otros tipos de ponderación que se han podido ver en el apartado anterior.

2.3.3. Error de seguimiento

Definición 2.12. *Llamamos **error de seguimiento** a la diferencia entre el rendimiento del portafolio y el índice que ese portafolio quiere replicar.*

Este tipo de error viene dado principalmente por las comisiones o costes de transacción que el índice no tiene en cuenta pero que el inversor, por el hecho de trabajar con ese portafolio, debe asumir. Hay dos tipos de variables que afectan a este tipo de error: exógenas (reajuste del índice o cambios en la composición de los activos) y endógenas (volatilidad del portafolio o cambios en el precio). Este tipo de error se puede cuantificar a través de la desviación típica o del error absoluto medio. El inversor puede sacar información relevante con esta variable: si su portafolio obtiene un rendimiento bajo y el error de seguimiento es alto, implica que el portafolio no es óptimo y debe realizar algún cambio.

2.3.4. Sesgo de eficiencia

Definición 2.13. Llamamos *sesgo de eficiencia* a la diferencia entre la rentabilidad del portafolio del mercado y la rentabilidad que ofrece la frontera eficiente dado el mismo nivel de riesgo.

En la práctica, como el mercado no se comporta de manera totalmente eficiente (y por eso existen estos sesgos que estamos estudiando), el portafolio del mercado no se encuentra en la frontera eficiente. El sesgo de eficiencia, pues, mide la diferencia que le separa de la frontera y pone de manifiesto la dificultad de encontrar portafolios o índices que se sitúen en ella. Es obvio que, como más eficiente sea el mercado, menor será el sesgo de eficiencia debido a que el portafolio del mercado se situará más próximo a la frontera eficiente.

2.3.5. Sesgo activo

Definición 2.14. Llamamos *sesgo activo* a la diferencia entre la rentabilidad del mercado y la rentabilidad de la cartera de inversión más eficiente.

La inversión activa tiene sentido cuando el mercado no es eficiente y los índices que representan al mercado tienen un alto volumen de sesgo, permitiendo pues al inversor obtener un rendimiento mayor al del mercado con una estrategia de inversión correcta. Si bien uno puede pensar que la inversión activa y la inversión pasiva son contrarias, se complementan hasta el punto que el sesgo activo se puede minimizar combinando ambas estrategias.

2.4. Propiedades

Un índice, para reducir el sesgo en su totalidad, debe cumplir, o intentar cumplir, las siguientes propiedades para ser considerado una buena alternativa de inversión:

- Representativo: Aproxima de forma fiel el comportamiento del mercado.
- Completo: Representa todos los activos posibles.
- Factible: Compuesto por activos negociables en el mercado.
- Ponderado por capitalización: El peso de los activos viene dado por la capitalización bursátil de cada uno.

2.5. Indicador de Herfindhal

Definición 2.15. Sea n el número de activos del índice y w_i el peso de cada uno de ellos, el indicador de Herfindhal mide el grado de concentración del índice tal que

$$\text{HHI} = \sum_{i=1}^n w_i^2.$$

Tal y como está definido, sabemos que $(\frac{1}{n} < \text{HHI} < 1)$. Como mayor sea el índice de Herfindhal, mayor será el grado de concentración y, consecuentemente, menor diversificación de los activos. De este modo, el valor mínimo del índice de Herfindhal se alcanza cuando todos los pesos son iguales.

2.6. Ratios financieras

Las ratios son herramientas financieras muy importantes que nos permiten tener una visión global de una compañía y saber el estado en el que se encuentra. Al tratarse de variables numéricas, permiten al inversor poder analizar una empresa de una forma relativamente cómoda y poder comparar compañías con características similares. Existen muchos tipos y es muy importante tener en cuenta el mayor número de ratios posible, ya que se tienen que analizar en conjunto y estudiar el origen de posibles anomalías.

Definición 2.16. La ratio **PER** (o *price-to-earning ratio*) determina la cantidad que los inversores están dispuestos a pagar en relación con los beneficios tal que:

$$PER = \frac{\text{Cotización de la acción}}{\text{Beneficio neto por acción}}.$$

Este indicador es muy útil para saber los años que tardará el inversor en recuperar la inversión inicial o las veces que el inversor paga el beneficio neto al comprar una acción y se utiliza para comparar compañías dentro de una misma industria.

Definición 2.17. La ratio **PSR** (o *price-to-sales ratio*) determina la valoración en el mercado de una compañía por cada unidad monetaria de sus ventas tal que:

$$PSR = \frac{\text{Capitalización bursátil}}{\text{Ingresos}}.$$

Este indicador, a diferencia del PER, se mantiene estable ante cambios en el gasto de la compañía, al tener en cuenta las ventas de la empresa en vez de los beneficios, pero a su vez no tiene en cuenta el nivel de deuda. Es muy útil para encontrar empresas con baja capitalización bursátil a un precio atractivo, mientras que no lo es tanto para empresas grandes ya que el valor suele ser muy alto.

Definición 2.18. La ratio **PBR** (o *price-to-book ratio*) determinada la valoración que tiene una compañía en el mercado en función de sus resultados contables tal que:

$$PBR = \frac{\text{Capitalización bursátil}}{\text{Valor neto contable}}.$$

El valor neto contable es el resultado de sustraer los activos intangibles y el pasivo del total del activo. Esta medida no es muy útil pues en empresas tecnológicas cuya propiedad intelectual es muy importante, ya que aquí no queda reflejado.

Definición 2.19. La *ratio sharpe* determina el exceso de rendimiento de un portafolio por cada unidad de riesgo tal que:

$$S_a = \frac{r_p - r_f}{\sigma_p}.$$

El concepto r_f o tasa libre de interés hace referencia al rendimiento de aquella inversión que no tiene ningún tipo de riesgo. En la práctica, se usa el tipo de interés ofrecido por los bonos del tesoro, entre otros. Una ratio sharpe alta indica un buen rendimiento del portafolio en función de su volatilidad, con lo que a priori, como más alto sea este valor más interesante es el portafolio en cuestión. Esta ratio es muy útil a la hora de realizar modificaciones en la composición de una cartera ya que compara las dos variables más importantes para su valoración: la incorporación de un activo puede bajar el rendimiento global de la cartera, pero si éste reduce la volatilidad de mayor modo el portafolio puede ser incluso mejor.

2.7. Exchange Trade Funds (ETF's)

En el mercado de derivados, aquel cuyos productos financieros varían el precio en función del subyacente o activo de referencia, los índices son muy utilizados como subyacente de productos derivados, referenciando así el rendimiento de un activo a la evolución del índice. A todo esto, la alternativa por excelencia que existe para invertir en los índices bursátiles son los ETF's.

Definición 2.20. *Los **ETF's** o **exchange trade funds** son valores o fondos que cotizan en el mercado de valores y que replican el comportamiento de un índice, un sector u otro tipo de activos.*

Puede haber muchos ETF's que repliquen el mismo índice. Sin ir más lejos, que imiten el comportamiento del S&P 500 hay: Vanguard S&P 500 ETF (VOO), SPDR S&P 500 ETF (SPY), iShares Core S&P 500 Index Fund ETF (IVV), y muchos más. Y aunque todos traten de replicar el mismo índice, sus características pueden ser ligeramente distintas. La principal diferencia es el grado de apalancamiento de cada ETF (multiplican por un coeficiente el rendimiento del índice), pues utilizan derivados financieros, como opciones y futuros, para tratar de obtener un mayor rendimiento que el propio índice. Esta práctica aumenta el riesgo del fondo, con lo que en caso de pérdidas, estas son más grandes.

De este modo, los ETF's son la herramienta financiera principal y más simple para invertir en un producto que replica el comportamiento de un índice, y poder beneficiarse de los movimientos del mercado. A este tipo de inversión se le llama gestión pasiva. Las dos principales ventajas de la gestión pasiva y los ETF's respecto la gestión activa y los fondos de inversión son:

1. La comisión que conlleva invertir en ETF's es mucho menor que la que conlleva invertir en fondos de inversión, ya que la gestión pasiva no requiere del estudio de oportunidades del mercado ni la manipulación continua de los activos en cartera que sí requiere la gestión activa.
2. Mientras que los ETF's se pueden negociar y realizar órdenes de compra y venta durante todo el día que el mercado está abierto (como si de acciones se trataran), los fondos de inversión solo permiten operaciones al cierre del día de mercado.
3. Existe la posibilidad de negociar a la baja a través de los ETF's inversos, pues si el inversor cree que un mercado esta en declive puede obtener rendimiento con ello.

3. S&P 500

Es el índice por excelencia de Estados Unidos, compuesto por las 500 empresas con mayor capitalización bursátil que cotizan en *NYSE* o *NASDAQ* (los dos mercados de valores más importantes del país). Se crea el año 1957 y es uno de los índices más antiguos que existen a día de hoy. El número exacto de valores que componen el índice puede ser ligeramente diferente a 500, ya que hay empresas con diferentes tipos de acciones que se computan por separado. En la actualidad hay 505 valores, y la capitalización bursátil de cada compañía se calcula sobre el capital flotante.

El S&P 500 representa de forma rigurosa el comportamiento del mercado de EEUU (al menos hasta ahora), pero su composición ha cambiado notablemente a lo largo de los años. Un dato significativo es la evolución del índice de Herfindhal en los últimos 30 años:

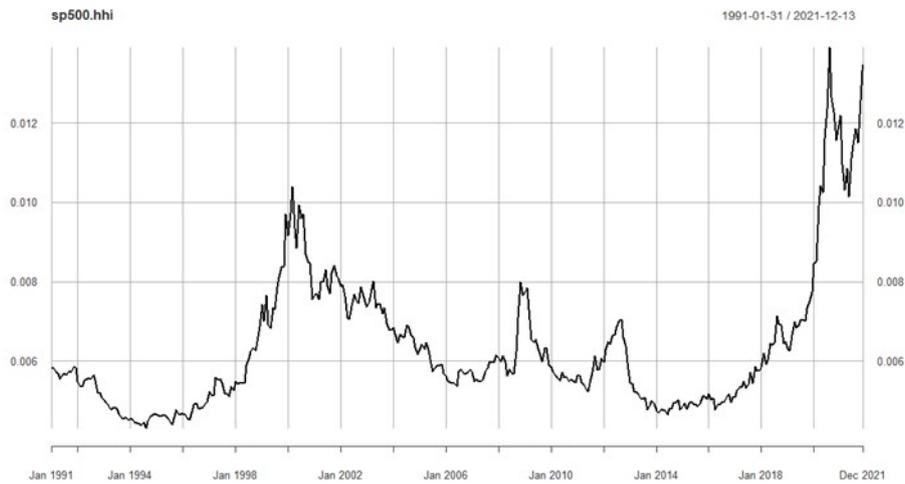


Figura 3.1 (Fuente: Bloomberg and Saxo Group)

Los máximos históricos alcanzados estos últimos años es consecuencia de la presencia de empresas y sectores que acaparan buena parte de la capitalización total del índice, provocando una mayor concentración de compañías y una menor diversificación de los activos. Esto no implica que el índice sea mejor o peor, ya que la representatividad del mercado puede ser lograda precisamente con estos niveles de concentración (o no).

3.1. Requisitos

El organismo que controla todos las modificaciones o comportamientos del S&P 500 es el *S&P Index Committee*, y son los que tienen la última palabra en cuanto a la exclusión o inclusión de cualquier compañía en el índice. Son los encargados de analizar exhaustivamente cada una de las empresas que componen el índice y, en caso de detectar cualquier irregularidad o prever cualquier anomalía que pueda afectar a la estructura del índice, tomar las medidas oportunas en cada momento. Los requisitos generales para que una compañía pueda entrar a formar parte del S&P 500 son:

1. Operar en Estados Unidos.
2. Tener un mínimo de capitalización bursátil de \$13 mil millones, salvo excepciones.
3. Haber obtenido beneficios en el último trimestre y en el último año.
4. Alta liquidez de sus acciones, eso implica haber negociado un número mayor de acciones que las que tiene en circulación el último año y haber negociado al menos 250.000 acciones en cada uno de los últimos seis meses.
5. Tener un mínimo de 50 % de las acciones accesibles o negociables para el público.

3.2. Propiedades

Algunas de las propiedades que posee el S&P 500 son las siguientes:

1. Muy sensible a los cambios del mercado debido a la ponderación por capitalización bursátil de los activos.
2. Representa el 80 % de la capitalización bursátil de Estados Unidos.
3. Distingue empresas de hasta 11 sectores dentro del mercado: tecnología, comunicación, inmobiliario, financiero, consumo básico, salud, energía, materiales, consumo discrecional, industrial y utilidades.
4. Se actualiza diariamente.
5. Sus activos pueden ser de dos tipos: acciones o REIT's (real estate investment trust), que son compañías que se dedican al sector inmobiliario y su principal fuente de ingreso proviene de generar rentas a través de estos inmuebles. Se caracterizan por tener un alto reparto de dividendos y una baja fiscalidad.

3.3. Sectores

Una de las claves por las que el S&P 500 se dice que es un buen indicador del mercado en EEUU es porque está compuesto por muchas compañías de muchos sectores. Esto le permite poder representar de manera fiel las fluctuaciones del mercado. El grado de representación que tiene cada sector dentro del índice es el siguiente:

Sector	Peso	nº compañías
Tecnología	29,17 %	75
Comunicación	10,16 %	27
Inmobiliario	2,77 %	29
Financiero	10,69 %	65
Consumo básico	5,88 %	32
Salud	13,29 %	64
Energía	2,67 %	21
Materiales	2,56 %	28
Consumo discrecional	12,54 %	63
Industrial	7,77 %	73
Utilidades	2,50 %	28

El claro sector que goza de mayor representación en el índice es el tecnológico. Y esto no solo se debe por poseer el mayor número de compañías dentro del índice sino porque son también las que tienen una mayor capitalización bursátil: Apple y Microsoft entre las más destacadas. Por otra parte, se puede observar que hay cuatro sectores (utilidades, materiales, inmobiliario y energía) cuya representación no excede el 3 %.

Parte II

Estudio de la técnica de los mapas autoorganizados

En la parte II de la memoria explicamos en qué consisten los mapas autoorganizados (en inglés, SOM o self-organizing maps), una técnica del lenguaje automatizado (en inglés, machine learning). Pondremos en contexto el lenguaje automatizado, presentaremos los diferentes tipos de aprendizaje que existen, y entraremos en detalle en qué consiste esta técnica: bajo qué supuestos se utiliza, con qué algoritmo trabaja, qué propiedades tiene, cuáles son sus ventajas y desventajas, etc.

4. Mapas autoorganizados

Para comprender bien que son los mapas autoorganizados y como se comportan, es necesario presentar el lenguaje automatizado y los tipos de aprendizaje que existen.

4.1. Introducción

Definición 4.1. *El lenguaje automatizado (o machine learning en inglés) es una rama de la ciencia computacional que utiliza herramientas estadísticas cuyo objetivo es desarrollar técnicas o algoritmos que permitan aprender a las computadoras.*

Definición 4.2. *Llamamos **red neuronal** al modelo de lenguaje automatizado compuesto por neuronas conectadas entre sí, que reciben información de los datos de entrada y se ajustan para producir unos valores de salida finales.*

4.2. Modelos de aprendizaje

Existen dos modelos de aprendizaje diferentes en el lenguaje automatizado: modelo de aprendizaje supervisado y modelo de aprendizaje no supervisado.

Definición 4.3. *El **modelo supervisado** es un tipo de aprendizaje cuyo algoritmo entrena los datos de entrada, previamente etiquetados por un agente externo, para obtener un resultado determinado.*

Este tipo de modelos disponen de una salida objetivo asociada a cada patrón de entrada y requieren de un agente externo para controlar el correcto funcionamiento de la red. A medida que los datos de entrada se introducen en el modelo, éste modifica sus conexiones hasta quedar completamente ajustado. De este modo, el modelo ha encontrado una función que es capaz de predecir el valor correspondiente a datos de entrada no vistos anteriormente.

Definición 4.4. *El **modelo no supervisado** es un tipo de aprendizaje cuyo algoritmo clasifica la información de entrada en función de patrones o características similares sin tener previamente una salida objetivo.*

En este tipo de modelos, la red descubre por si misma rasgos comunes, regularidades, correlaciones o categorías entre los datos de entrada que luego incorpora en su estructura interna de conexiones. No necesita de la supervisión de un agente externo, ya que es el propio modelo quien detecta los patrones entre los datos de entrada. El modelo es muy utilizado para la agrupación o de datos o clustering.

De este modo, la principal diferencia entre ambos modelos de aprendizaje es que uno dispone de una salida objetivo en la que base su entrenamiento y poder así aplicar el mismo proceso de entrenamiento para nuevos datos introducidos en el sistema mientras que el otro lo que hace es descubrir características similares entre los datos de entrada para obtener un resultado determinado.

4.3. Definición

Definición 4.5. *Los **mapas autoorganizados** son redes neuronales artificiales que, a través de un aprendizaje competitivo y no supervisado, transforman datos de entrada de alta dimensión (inputs) en datos de salida de baja dimensión (outputs) que se representan, por norma general, en mapas 2D.*

Esta técnica fue descubierta por Teuvo Kohonen, alrededor del 1980, y ha sido objeto de estudio por múltiples autores (ver [7] y [8]) desde entonces. Su objetivo es descubrir características o patrones similares entre un conjunto de datos de entrada y representar los resultados en un mapa de salida que permita visualizar fácilmente relaciones relevantes entre los datos.

Los mapas autoorganizados están formados por neuronas, que se caracterizan por su vector de pesos asociado, y son las encargadas de conectar \mathbb{R}^M con \mathbb{R}^2 , siendo M la dimensión de los datos de entrada (o número de variables). El método de entrenamiento, que veremos más adelante, se basa en la actualización de los pesos para que estos sean altamente representativos de las agrupaciones formadas en cada neurona.

4.4. Estructura

La estructura de un mapa autoorganizado viene determinada por dos factores:

1. Tipología: El mapa puede ser de tipo rectangular o tipo hexagonal, y caracteriza las relaciones entre vecindades. En un mapa rectangular cualquier neurona tiene, como mucho, cuatro neuronas a su alrededor o entorno más cercano. En cambio, en un mapa de tipo hexagonal, una neurona puede llegar a tener seis neuronas a su alrededor, con lo que el proceso de actualización involucra más neuronas. Esta diferencia queda ilustrada en la siguiente imagen:

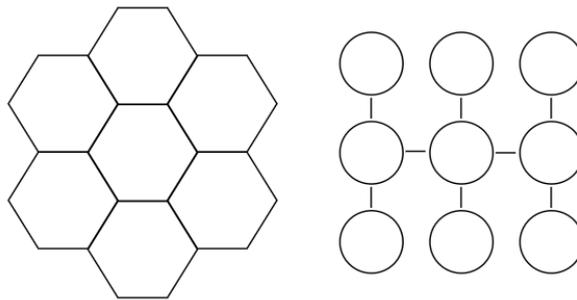


Figura 5.1.1 (Elaboración Propia)

2. Dimensión: El mapa, que se suele representar como una matriz cuadrada, puede tener cualquier cantidad de neuronas (5×5 , 10×10 , 10×5 ...) en función de las necesidades de cada estudio.

Matriz de pesos de cada neurona

Sea K el número de neuronas y M el número de variables que describen a cada dato de entrada, obtenemos la siguiente matriz de pesos W de dimensión $K \times M$:

$$W = \begin{matrix} w^{(1)} \rightarrow \\ w^{(2)} \rightarrow \\ w^{(3)} \rightarrow \\ \vdots \\ w^{(k)} \rightarrow \\ \vdots \\ w^{(K)} \rightarrow \end{matrix} \begin{pmatrix} w_{11} & w_{12} & w_{13} & \dots & w_{1m} & \dots & w_{1M} \\ w_{21} & w_{22} & w_{23} & \dots & w_{2m} & \dots & w_{2M} \\ w_{31} & w_{32} & w_{33} & \dots & w_{3m} & \dots & w_{3M} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ w_{k1} & w_{k2} & w_{k3} & \dots & w_{km} & \dots & w_{kM} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ w_{K1} & w_{K2} & w_{K3} & \dots & w_{Km} & \dots & w_{KM} \end{pmatrix} \quad (4.2)$$

Esta matriz de pesos, cuyas filas describen a cada neurona del mapa autoorganizado, se puede inicializar de múltiples maneras (ver más adelante) y es la que acaba determinando el resultado del mapa.

4.6. Algoritmo de aprendizaje

Esta sección explica el proceso que sigue el mapa autoorganizado cuyo algoritmo de aprendizaje se explica de la siguiente manera:

```
inicializacion del vector de pesos de cada neuronas;
for  $t = 1$  to  $t_{max}$  do
     $x \leftarrow$  eleccion del vector de entrada;
     $BMU \leftarrow$  neurona mas similar a  $x$ ;
    actualizacion del vector de pesos de la BMU y neuronas vecinas;
end
```

4.6.1. Inicialización

En la primera fase del algoritmo se asocia un vector de pesos $w^{(k)}$ a todas y cada una de las neuronas que forman el mapa autoorganizado, completando pues todos los elementos de la matriz de pesos W . Existen diferentes métodos para llevar a cabo esta inicialización:

1. Aleatoriamente: Dar valores aleatorios a los vectores de pesos $w^{(k)}$.
2. Según los datos de entrada: Consiste en asociar valores de los datos de entrada a los pesos de las neuronas de modo que la distribución inicial de ambas matrices sea similar.
3. Algoritmo de máxima distancia: Trata de asociar valores a las neuronas de modo que sus pesos estén situados a la mayor distancia posible.

Si bien es cierto que el método de inicialización de las neuronas no debe afectar al resultado final del mapa, sí que puede ayudar a una convergencia hacia la solución más rápida. En aquellos casos en que el número de iteraciones máximo es elevado, el método utilizado es irrelevante. Por simplicidad se suele usar la inicialización aleatoria.

4.6.2. Elección

Esta fase del algoritmo es trivial y consiste solamente en elegir aleatoriamente un vector de entrada $x^{(n)}$ que permanece estable durante toda la iteración. Al finalizar la misma, se procede a la elección aleatoria de otro vector de entrada $x^{(n)}$ para la siguiente iteración.

4.6.3. Competición

En la fase de competición, las neuronas compiten entre ellas y solo una resulta vencedora. Para determinar la neurona vencedora, es necesario precisar antes la distancia que determina el resultado y que muy habitualmente es la distancia euclidiana entre dos vectores:

$$\|x^{(n)} - w^{(k)}\| = \sqrt{\sum_{i=1}^M (x_i^{(n)} - w_i^{(k)})^2}, \quad \forall k = 1, \dots, K.$$

Definición 4.6. Llamamos **BMU** (del inglés *best matching unit*) o **neurona vencedora** a aquella cuyo vector de pesos se encuentra a distancia mínima del vector de entrada $x^{(n)}$ y, por lo tanto, más similar. De este modo:

$$BMU = \underset{k}{\operatorname{argmin}} \|x^{(n)} - w^{(k)}\|.$$

4.6.4. Actualización

La actualización de las neuronas (ver Wilmott, Paul [9]) es la fase más compleja e importante del algoritmo, pues establece las propiedades de los mapas autoorganizados. Es necesario definir la distancia euclidiana entre dos neuronas $a = (a_x, a_y)$ y $b = (b_x, b_y)$ o puntos en un mapa 2D (ver figura 5.3.1):

$$\|(a, b)\| = \sqrt{(b_x - a_x)^2 + (b_y - a_y)^2}.$$

Definida la distancia, el proceso de actualización de las neuronas $\forall k = 1, \dots, K$ es

$$w^{(k)} \leftarrow w^{(k)} + \underbrace{\varepsilon_0 \cdot \exp\left(-\frac{t}{t_{max}}\right)}_{\alpha(t)} \cdot \underbrace{\exp\left(-\frac{\|(BMU, k)\|^2}{2 \cdot (\sigma_0)^2 \cdot \exp\left(-\frac{2t}{t_{max}}\right)}\right)}_{h(a,b)} \cdot (x^{(n)} - w^{(k)}),$$

donde ε_0 y σ_0 son parámetros. En el caso de tener estandarizadas las variables según distribución normal, se estima que sus valores correctos son $\varepsilon_0 \approx 0,2$ y $\sigma_0 \approx \frac{1}{2}\sqrt{K}$.

La función $\alpha(t)$, también llamada tasa de aprendizaje, se sitúa entre $0 < \alpha(t) < 1$ y es decreciente en el tiempo, es decir, la actualización es menor con el paso de las iteraciones ($\alpha(t+1) < \alpha(t)$). Esto evita oscilaciones grandes en las iteraciones finales que podrían afectar significativamente en el resultado del mapa.

La función $h(a, b)$, también llamada función de vecindad, es proporcional a la distancia entre las neuronas y controla el grado de actualización de las mismas. Debe cumplir:

$$h(a, b) \rightarrow 0 \quad \text{si} \quad \|(a, b)\| \rightarrow \infty$$

Esto provoca una actualización gradual de las neuronas: cuanto menor sea la distancia entre la neurona en cuestión y la BMU, mayor será la actualización. De esta manera, se asegura una mayor similitud entre neuronas vecinas.

4.7. Propiedades

El algoritmo presenta las siguientes propiedades:

1. Lentitud de entrenamiento (sobretudo cuando trabajamos con una gran cantidad de datos de entrada).
2. Si hay muchas variables, la visualización e interpretación de los resultados es complicada.
3. Principio de proximidad: dos neuronas vecinas en el espacio de salida también lo son en el espacio de entrada \mathbb{R}^M .
4. Sea $c = \underset{k}{\operatorname{argmin}} \|x^{(n)} - w^{(k)}\|$, se puede medir el error del método a través de

$$E_r = \sum_{i=1}^N \|x^{(i)} - w^{(c)}\|^2$$

Parte III

Construcción de una cartera óptima

Esta parte tiene como objetivo principal crear una cartera de inversión a partir de los activos que componen el S&P 500, y que lo optimice. Para ello utilizaremos una base de datos con información y ratios de todas las empresas que componen el índice. Todos estos datos serán tratados a través de RStudio y, utilizando unos paquetes predeterminados, crearemos el mapa autoorganizado que buscamos para tener agrupaciones de empresas en función de sus similitudes. De este modo, podremos escoger las empresas que mejor puedan encajar en nuestra cartera y comparar el rendimiento global de la misma con el del S&P 500.

5. Introducción

La creación de carteras de inversión se lleva haciendo desde hace muchos años y cada día existen más herramientas que permiten al inversor crear cualquier cartera que desee en función de su perfil. Sin ir más lejos, podemos entender los fondos como carteras de inversión, con unas regulaciones e intereses determinados.

El primer paso para la creación de la cartera es escoger los activos que la componen. Esta elección la haremos a través de la creación de un mapa autoorganizado con el programa RStudio, en función de la base de datos expuesta más adelante, y que nos agrupará todos los activos del S&P 500 en diferentes grupos. De todos ellos, seleccionaremos los grupos relevantes o más distintos para que la elección del activo sea coherente con un cierto grado de diversificación.

El siguiente paso es proponer tres carteras diferentes en función del método de ponderación escogido: pesos iguales, por capitalización bursátil y portafolio de mínima varianza.

Finalmente, a través de la herramienta gratuita online *Portfolio Visualizer*² accesible para todo el mundo, podremos analizar y valorar las tres carteras y compararlas con el comportamiento del S&P 500, teniendo en cuenta:

1. Espacio temporal: *2008-2021*. Se considera robusto realizar un análisis del comportamiento de las acciones en un periodo de 14 años con periodos anómalos dentro de él: crisis del 2008 y pandemia del 2020.
2. Índice de referencia: *Vanguard 500 Index Investor*. Es el índice que replica el comportamiento del S&P 500 y que, a su vez, describe el comportamiento del mercado de valores (en Estados Unidos).

6. Construcción de las carteras

6.1. Base de datos

Si bien es cierto que existe mucha información acerca del S&P 500, no es tan fácil encontrar información individualizada y relevante de cada compañía que compone el índice.

La base de datos utilizada (<https://stockmarketmba.com/stocksinthesp500.php>) está formada por las 505 empresas que componen el S&P 500 y contiene las siguientes variables:

1. Sector en el que pertenece la empresa.
2. Capitalización bursátil.
3. Dividendos anuales que reparten al accionista.
4. Ratio PER.
5. Ratio PSR.
6. Ratio PBR.

²<https://www.portfoliovisualizer.com/backtest-portfolio>

6.2. Creación del mapa autoorganizado

Este es el código en RStudio utilizado para la creación del mapa autoorganizado:

```
1 #Libreria necesaria para crear el mapa autoorganizado
2 library (kohonen)
3
4 #Llamo al fichero de datos donde tengo todas las variables
5 data ← read.csv(file.choose(), header = T)
6
7 #Estandarizo todas las variables (media=0 y desv_tip=1)
8 X ← scale(data[,-1])
9
10 #Inicializo los vectores de pesos del mapa de forma aleatoria
11 set.seed(2021)
12
13 #Defino la dimension del mapa autoorganizado
14 g ← somgrid(xdim=5, ydim=5, topo = "hexagonal")
15
16 #Creo el mapa autoorganizado
17 map ← som(X, grid=g, alpha=c(0.05, 0.01))
18
19 #Frecuencia de mapeo
20 plot(map, type="mapping", shape = "straight")
21
22 #Distancia entre vecindades
23 plot(map, type="dist.neighbours", shape = "straight")
24
25 #Distancia entre acciones del mismo neurona y su vector de pesos
26 plot(map, type = "quality", shape = "straight")
27
28 #Descubro a que neurona en particular pertenece cada accion
29 map$unit.classif
30
31 #Desgloso las acciones que se encuentran en cada neurona del mapa
32 which(map$unit.classif == 1)
33 which(map$unit.classif == 2)
34 ...
35 which(map$unit.classif == 25)
```

6.3. Selección de activos

A través del código desarrollado en el apartado anterior, podemos visualizar:

1. Frecuencia de mapeo: Este mapa nos proporciona la cantidad de acciones que contiene cada neurona del mapa sin identificar a cada una de ellas.

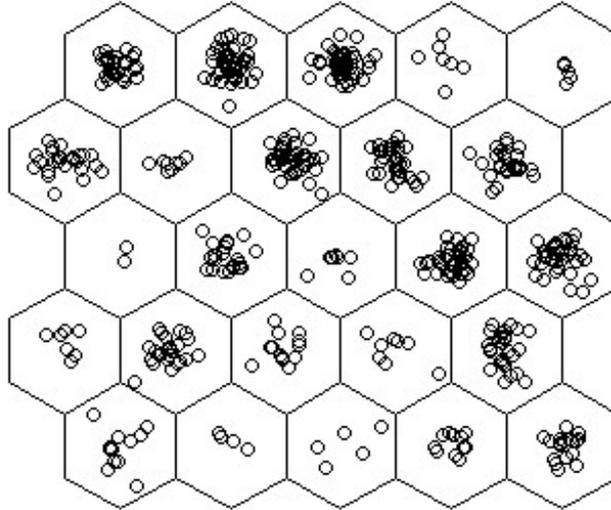


Figura 8.3.1

2. Distancia entre vecindades: Este mapa representa la similitud entre vecindades. Según la leyenda que aparece junto al mapa, las vecindades con colores rojos son aquellas que tienen distancias pequeñas y, consecuentemente, un grado de similitud elevado. En cambio, aquellas vecindades más claras indican que se encuentran a una distancia elevada de las otras y, consecuentemente, un grado de similitud bajo.

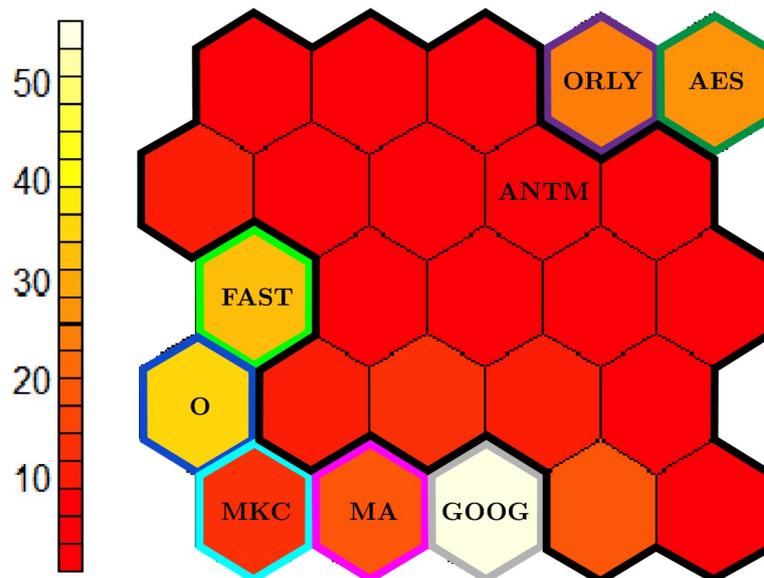


Figura 8.3.2

Respecto a la representación de este último mapa, hay que tener en cuenta:

1. La agrupación de neuronas se ha hecho en función de las distancias en el mapa. El número de grupos escogido podría haber sido uno u otro, pensando que el elegido es una buena manera de construir una cartera diversificada y poder analizar profundamente la relación entre los activos.
2. La selección del activo dentro de cada neurona se ha hecho en función del comportamiento individual de cada uno de ellos sobre el rendimiento general de la cartera, teniendo en cuenta factores como la varianza o la correlación con el mercado.

En la siguiente tabla podemos observar los ocho activos que hemos seleccionado junto con dos variables relevantes: sector donde pertenece cada acción y correlación con el mercado.

Nombre	Símbolo	Sector	Correlación S&P 500
Mastercard Incorporated	MA	Financiero	0.61
Alphabet Inc.	GOOG	Comunicación	0.61
Realty Income Corporation	O	Inmobiliario	0.48
Anthem, Inc.	ANTM	Salud	0.53
McCormick & Company	MKC	Consumo Básico	0.42
Fastenal Company	FAST	Industrial	0.62
The AES Corporation	AES	Utilidades	0.63
O'Reilly Automotive, Inc.	ORLY	Consumo Discrecional	0.42

Podemos observar que, para garantizar una cierta diversificación de los activos, hemos escogido acciones que pertenezcan a diferentes sectores, dejando sin representación los sectores de energía, materiales y tecnología. También podemos ver que la correlación que hay entre el rendimiento de los activos y del S&P 500 se mueve entorno el 40 % y 60 %, mientras que la correlación que existe entre los activos es la siguiente:

	MA	GOOG	O	ANTM	MKC	FAST	AES	ORLY
MA	-	0.50	0.24	0.27	0.26	0.35	0.31	0.40
GOOG	0.50	-	0.27	0.30	0.23	0.26	0.39	0.25
O	0.24	0.27	-	0.22	0.36	0.36	0.32	0.44
ANTM	0.27	0.30	0.22	-	0.24	0.29	0.34	0.26
MKC	0.26	0.23	0.36	0.24	-	0.24	0.32	0.41
FAST	0.35	0.26	0.36	0.29	0.24	-	0.23	0.40
AES	0.31	0.39	0.32	0.34	0.32	0.23	-	0.12
ORLY	0.40	0.25	0.44	0.26	0.41	0.40	0.12	-

Se puede observar claramente que la correlación máxima entre activos es de un 50 % entre Mastercard y Google, mientras que todas las otras correlaciones se mueven entorno el 20 % y 40 %, considerando pues un nivel aceptable de diversificación en la cartera. Dado el número de activos que tenemos, esta variable es muy importante ya que queremos evitar compañías cuyo comportamiento sea parecido y la correlación sea muy alta.

6.4. Constitución de las carteras

Siguiendo las hipótesis del problema de Markowitz, hemos determinado la rentabilidad mensual esperada (\bar{r}) y la desviación típica mensual (σ) de cada uno de los ocho activos que componen la cartera a partir de la media de las rentabilidades mensuales en el periodo establecido, construyendo pues la siguiente tabla:

	MA	GOOG	O	ANTM	MKC	FAST	AES	ORLY
\bar{r}	1,93 %	1,69 %	0,87 %	1,36 %	1,10 %	1,36 %	0,50 %	2,10 %
σ	7,47 %	7,63 %	6,64 %	8,34 %	4,88 %	7,66 %	8,68 %	6,86 %

Para calcular la matriz de varianzas y covarianzas solo es necesario la desviación típica de cada activo calculada en la tabla anterior y la matriz de correlaciones entre los activos, que está expuesta más adelante. De todos modos, a través del programa Excel la podemos calcular directamente obteniendo el siguiente resultado:

	MA	GOOG	O	ANTM	MKC	FAST	AES	ORLY
MA	0,56 %	0,28 %	0,11 %	0,16 %	0,08 %	0,20 %	0,20 %	0,20 %
GOOG	0,28 %	0,58 %	0,13 %	0,18 %	0,07 %	0,15 %	0,25 %	0,12 %
O	0,11 %	0,13 %	0,44 %	0,11 %	0,11 %	0,18 %	0,17 %	0,19 %
ANTM	0,16 %	0,18 %	0,11 %	0,69 %	0,08 %	0,18 %	0,24 %	0,14 %
MKC	0,08 %	0,07 %	0,11 %	0,08 %	0,24 %	0,08 %	0,13 %	0,12 %
FAST	0,20 %	0,15 %	0,18 %	0,18 %	0,08 %	0,58 %	0,15 %	0,20 %
AES	0,20 %	0,25 %	0,17 %	0,24 %	0,13 %	0,15 %	0,75 %	0,06 %
ORLY	0,20 %	0,12 %	0,19 %	0,14 %	0,12 %	0,20 %	0,06 %	0,47 %

Con todo esto, es posible construir el portafolio de mínima varianza a través de dos restricciones: minimizar la desviación típica del portafolio y que la suma de los pesos totales sea del 100 %. De este modo ya tenemos las tres carteras que queremos analizar:

	Eq. weighted	Cap. weighted	Min. var.
MA	12,50 %	17,50 %	6,99 %
GOOG	12,50 %	19,50 %	10,04 %
O	12,50 %	9,50 %	12,88 %
ANTM	12,50 %	15,50 %	8,02 %
MKC	12,50 %	7,50 %	48,82 %
FAST	12,50 %	11,50 %	7,37 %
AES	12,50 %	5,50 %	0,71 %
ORLY	12,50 %	13,50 %	5,17 %
\bar{r}	1,36 %	1,51 %	1,28 %
σ	4,48 %	4,67 %	4,02 %

Respecto al portafolio por capitalización bursátil, debido a la diferencia muy grande que había entre la capitalización de las compañías, hemos establecido un máximo y un mínimo para que todos los activos tengan cierta representación y establecido un 2 % de diferencia entre un activo y el siguiente en volumen de capitalización.

6.5. Representación de la frontera eficiente

Una vez obtenido el portafolio de mínima varianza, si encontramos otra solución del problema de Markowitz, podremos construir la frontera eficiente en su totalidad. Imponiendo que la rentabilidad sea máxima encontramos un segundo portafolio que se encuentra en la frontera eficiente, y que llamamos portafolio de máxima rentabilidad. A partir de estos dos, podemos encontrar portafolios intermedios fijando una rentabilidad determinada (que esté entre los dos portafolios extremos) y minimizando la desviación típica. Los portafolios eficientes que hemos encontrado son los siguientes:

	Min. var.	2	3	4	5	6	7	8	9	Max. \bar{r}
MA	6,99 %	8,75 %	10,55 %	12,35 %	15,09 %	17,83 %	20,35 %	22,83 %	26,30 %	0,00 %
GOOG	10,04 %	11,30 %	12,47 %	13,62 %	14,38 %	15,14 %	15,76 %	16,37 %	11,77 %	0,00 %
O	12,88 %	8,81 %	4,33 %	0,00 %	0,00 %	0,00 %	0,00 %	0,00 %	0,00 %	0,00 %
ANTM	8,02 %	7,91 %	7,58 %	7,26 %	6,52 %	5,79 %	4,61 %	3,37 %	0,00 %	0,00 %
MKC	48,82 %	45,62 %	42,05 %	38,32 %	30,66 %	23,00 %	14,27 %	5,35 %	0,00 %	0,00 %
FAST	7,37 %	6,67 %	5,86 %	5,02 %	2,68 %	0,34 %	0,00 %	0,00 %	0,00 %	0,00 %
AES	0,71 %	0,00 %	0,00 %	0,00 %	0,00 %	0,00 %	0,00 %	0,00 %	0,00 %	0,00 %
ORLY	5,17 %	10,95 %	17,16 %	23,42 %	30,67 %	37,91 %	45,01 %	52,08 %	61,93 %	100,00 %
\bar{r}	1,28 %	1,37 %	1,46 %	1,55 %	1,64 %	1,73 %	1,82 %	1,92 %	2,01 %	2,10 %
σ	4,02 %	4,05 %	4,12 %	4,25 %	4,43 %	4,67 %	4,96 %	5,30 %	5,70 %	6,83 %

Por último, y con todos estos portafolios, podemos representar ya la frontera eficiente:

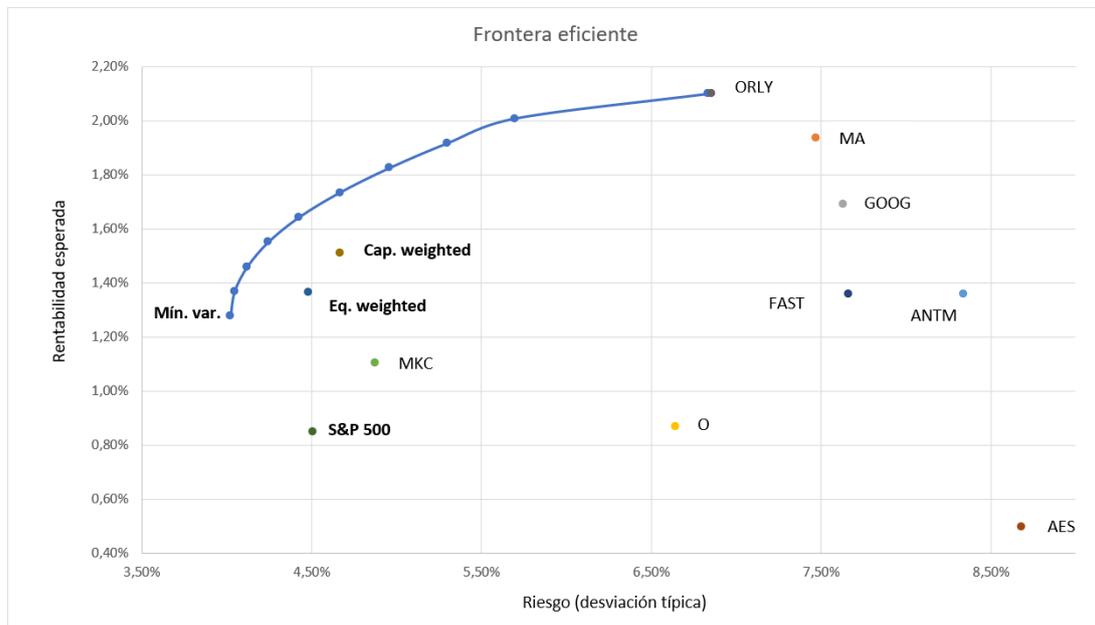


Figura 8.5.1

Esta información es muy importante para ver el grado de ineficiencia de las carteras construidas desde el punto de vista del inversor racional. La cartera por capitalización y la cartera con igualdad de pesos, si bien es cierto que no están en la frontera eficiente como era de esperar, se sitúan relativamente cerca de ella, con lo que la pérdida de eficiencia no es muy alta. La cartera que queda muy lejos de la frontera eficiente es el índice S&P 500, y se ve claramente que tiene un nivel de riesgo similar a la cartera con igualdad de pesos pero con una rentabilidad mucho menor.

7. Valoración de las carteras

A continuación podemos observar la representación gráfica del rendimiento de las tres carteras en comparación con el rendimiento del SP&500 en el periodo comprendido entre los años 2008-2021. Tal y como consta en la leyenda, el portafolio con pesos iguales es representado con color azul, por capitalización en color rojo, el de mínima varianza con color amarillo y el S&P 500 con color verde.

Portfolio	Initial Balance	Final Balance	CAGR	Stddev	Best Year	Worst Year	Max. Drawdown	Sharpe Ratio	Sortino Ratio
Eq. weighted	\$10,000	\$76,086	15.60%	15.71%	45.67%	-30.33%	-36.06%	1.07	1.77
Cap. weighted	\$10,000	\$95,345	17.48%	16.37%	52.17%	-32.11%	-36.86%	1.11	1.89
Min. var.	\$10,000	\$68,515	14.74%	14.46%	33.08%	-21.56%	-28.38%	1.12	1.86
Vanguard 500 Index Investor	\$10,000	\$32,546	8.79%	15.67%	32.16%	-37.03%	-49.69%	0.70	1.04

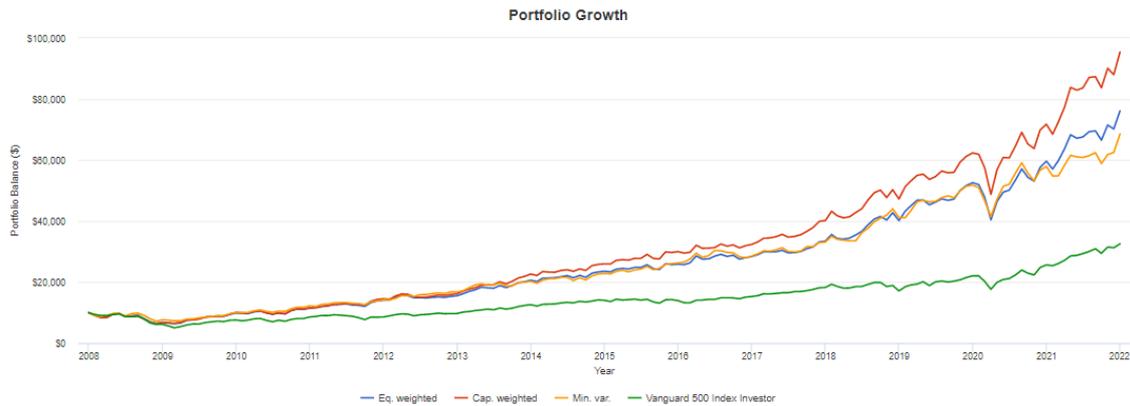


Figura 9.1

Si bien es cierto que las tres carteras que hemos creado se sitúan en valores superiores a la del índice a lo largo de todo el periodo, se puede observar que reflejan de manera bastante fiel todas las anomalías bruscas que sufre el mercado, tanto en periodos alcistas como bajistas. En cuanto al análisis numérico y comparación entre las carteras y el S&P 500, nos basaremos en las cuatro variables que aparecen en la siguiente tabla:

	Eq. weighted	Cap. weighted	Min. var.	S&P 500
CAGR	15,60 %	17,48 %	14,74 %	8,79 %
Desviación Estándar	15,71 %	16,37 %	14,46 %	15,67 %
Ratio Sharpe	1,07	1,11	1,12	0,70
Correlación Mercado EEUU	0,88	0,86	0,75	1,00

1. CAGR: Si bien es cierto que el rendimiento anualizado de las tres carteras que hemos creado se sitúa en buenos niveles de rentabilidad oscilando entorno el 15 %, el portafolio ponderado por capitalización bursátil destaca ligeramente respecto los otros llegando a alcanzar 17,48 %. Por otra parte, el 8,79 % correspondiente al rendimiento anualizado del S&P 500, no destaca como las otras carteras aunque lo podemos considerar un nivel aceptable de rentabilidad. Esta diferencia entre el S&P 500 y las otras tres carteras, a largo plazo resulta una diferencia abismal para el inversor, y que en ningún caso elegiría el índice como alternativa de inversión.

2. Desviación típica: Con un 14,74 %, el portafolio de mínima varianza es, obviamente, el que minimiza la desviación típica (anualizada). Después tenemos tanto el S&P 500 como la cartera con igualdad de pesos que se sitúan entorno el 15,70 %. Finalmente, la cartera que presenta mayor riesgo es la ponderada por capitalización llegando hasta el 16,37 %. Mientras que en las tres carteras esta variable es directamente proporcional al CAGR visto en el punto anterior (mayor riesgo en aquellas con mayor rentabilidad), el S&P 500 tiene la menor rentabilidad pero desviación típica intermedia.
3. Ratio Sharpe: Este indicador pone de manifiesto que el S&P 500 no dispone de una rentabilidad suficiente por el riesgo total que comporta ya que su ratio sharpe se mueve entorno al 0,7, pues indica que por cada unidad de riesgo solo rentabiliza una séptima parte. En cambio, las otras carteras cuya ratio sharpe se sitúa por encima de 1 refleja la obtención de una mayor rentabilidad por unidad de riesgo. De todos modos, tanto el portafolio por capitalización como el de mínima varianza disponen de una ratio sharpe ligeramente superior al portafolio con igualdad de pesos.
4. Correlación con el mercado de EEUU: Este indicador, aunque no sea vinculante para el inversor, proporciona información muy interesante de cada cartera. Era de esperar que el S&P 500 tuviera un comportamiento idéntico o muy parecido al del mercado (podría ser ligeramente diferente en función del periodo). En cuanto a las tres carteras que hemos construido, el portafolio con igualdad de pesos (0,88) y el portafolio por capitalización (0,86) obtienen muy buen resultado, dejando en evidencia que replican fielmente el comportamiento del mercado (que ya podíamos intuir con la gráfica). Dicha variación se produce, en gran parte, por el mayor rendimiento de las tres carteras respecto al S&P 500, cosa que beneficia al inversor. También tenemos el portafolio de mínima varianza que, con un 0,75 de correlación con el mercado, es también representativo aunque no llega al nivel de las otras carteras.

8. Conclusiones

El objetivo de esta sección era crear diferentes carteras de activos que fueran más óptimas que el índice S&P 500 en cuanto a la relación riesgo-rentabilidad siempre priorizando unos retornos atractivos (por encima del índice).

Una vez elegidos los activos que iban a conformar nuestra cartera (a partir de la técnica SOM), hemos creado tres carteras con ponderaciones distintas para dichos activos. También hemos construido la frontera eficiente a través de nueve carteras auxiliares que cumplen las ecuaciones de Markowitz, situando en ella las tres carteras creadas y el S&P 500, estudio a partir del cual hemos podido observar claramente la mejor posición que presentan las ya mencionadas carteras.

De esta forma, concluir que los índices se deben tratar como lo que son: representaciones del mercado o de una parte de él, y aunque los inversores (especialmente los de gestión pasiva) vean los índices como fuentes de alta rentabilidad a largo plazo, estos pueden ofrecer retornos realmente inferiores y con mayor volatilidad que una cartera formada bajo criterios esenciales como los que hemos escogido en este caso.

Si algo quiero dejar claro es que invertir implica riesgo. Y lo más importante, actuaciones pasadas no garantizan resultados futuros. Es por todo esto que dejo en manos del lector la elección de la cartera que considere óptima en función de sus intereses y su perfil.

Referencias

- [1] Sigman, Karl. <http://www.columbia.edu/~ks20/FE-Notes/4700-07-Notes-portfolio-I.pdf>. “Portfolio mean and variance.” (2005).
- [2] Zivot, Eric. <https://faculty.washington.edu/ezivot/econ424/portfolioTheoryMatrix.pdf>. “Chapter 1: Portfolio Theory with Matrix Algebra.” (2013).
- [3] Sigman, Karl. <http://www.columbia.edu/~ks20/FE-Notes/4700-07-Notes-funds.pdf>. “Fund theorems.” (2005).
- [4] Qontigo.com. https://www.stoxx.com/document/Indices/Common/Indexguide/stoxx_calculation_guide.pdf. “Creating an Investment Intelligence Advantage.” (2021).
- [5] S&P Dow Jones Indices. <https://www.spglobal.com/spdji/en/documents/methodologies/methodology-index-math.pdf>. “Index Mathematics Methodology.” (2021).
- [6] Andreu Corbatón, Jordi. “Market Indices: Bases, Biases and Beyond.” Universitat de Barcelona. (2009).
- [7] Ponmalai, Ravi, and Chandrika Kamath. “Self-Organizing Maps and Their Applications to Data Analysis.” No. LLNL-TR-791165. Lawrence Livermore National Lab. (LLNL), Livermore, CA (United States). (2019).
- [8] Deboeck, Guido, and Teuvo Kohonen, eds. “Visual Explorations in Finance: with Self-Organizing Maps.” Springer Science & Business Media. (2013).
- [9] Wilmott, Paul. “Machine Learning: An Applied Mathematics Introduction.” Panda Ohana Publishing. (2019).

Apéndice A: Código en R de la Parte III

```
#Librería necesaria para crear el mapa autoorganizado
1 library (kohonen)

#Llamo al fichero de datos donde tengo todas las variables
1 data ← read.csv(file.choose(), header = T)

#Estandarizo todas las variables (media=0 y desvtip=1)
1 X ← scale(data[, -1])

#Inicializo los vectores de pesos del mapa autoorganizado de forma aleatoria
1 set.seed(2021)

#Defino la dimensión del mapa autoorganizado
1 g ← somgrid(xdim=5, ydim=5, topo = "hexagonal")

#Creo el mapa autoorganizado
1 map ← som(X, grid=g, alpha=c(0.05, 0.01))

#Dibujo el mapa con el nº de acciones que pertenecen a cada neurona
1 plot(map, type = "mapping", shape = "straight")

#Dibujo el mapa con el nº de acciones que pertenecen a cada neurona (en colores)
1 plot(map, type = "counts", shape = "straight")

#Dibujo el mapa con la distancia que existe entre vecindades
1 plot(map, type = "dist.neighbours", shape = "straight")

#Mapa que representa la distancia entre las acciones de cada neurona y su vector de pesos
1 plot(map, type = "quality", shape = "straight")

#Descubro a qué neurona en particular pertenece cada acción
1 map$unit.classif

[1] 3 3 3 3 3 6 9 2 9 9 2 11 9 9 9 9 9 2 2 4 17 8 8 5 17 8 24 4 14 8 4 8 23 8 4 4 10 8 8 4 5 4
19 8 8 4 12 14 7 14 8 24 7 8 5 8 8 14 21 5 18 18 18 18 19 15 15 14 23 16 12 6 20 22 22 4
19 1 1 7 23 14 16 5 7 25 15 19 18 14 18 18 15 22 6 21 10 22 23 18 14 14 15 18 5 12 19 12
15 10 23 18 16 18 24 17 7 18 7 10 20 19 21 14 18 13 21 18 21 6 19 18 19 16 15 18 12 7 21
19 5 20 23 23 7 23 14 7 7 7 15 23 16 23 14 17 1 21 7 7 15 12 24 16 21 7 23 7 12 16 23 13
16 14 21 20 16 18 18 24 10 21 14 12 14 18 23 18 20 23 14 19 10 10 21 10 16 16 10 23 18
19 5 23 22 10 23 15 10 7 7 18 23 14 14 20 20 21 16 24 17 12 16 5 7 21 23 18 25 20 11 20
23 18 7 16 23 7 23 22 22 12 19 16 5 18 22 19 19 16 12 15 14 14 12 18 19 15 22 20 21 5 22
```

```
19 20 19 23 14 14 15 5 15 17 19 15 5 10 1 23 7 16 22 22 12 12 5 6 16 1 23 18 23 21 23 10
14 23 20 13 17 18 20 1 22 16 13 14 23 25 22 22 18 12 21 17 1 22 14 1 16 23 20 18 21 17 5
12 16 15 14 10 21 18 10 18 15 23 15 10 16 22 10 21 7 13 21 23 13 1 15 15 23 19 19 23 18
21 22 10 22 18 14 22 15 4 19 23 14 21 21 20 20 22 16 15 1 7 12 23 10 21 18 15 22 6 10 21
4 16 23 14 19 22 22 23 24 22 22 15 14 22 12 18 25 15 10 20 21 10 20 15 12 21 23 20 22 22
15 14 23 7 18 19 1 21 16 15 19 5 15 10 14 18 23 21 23 23 15 20 19 14 15 12 10 5 22 22 23
25 20 20 19 22 18 18 22 19 22 14 23 15 19 22 20 15 22 14 12 22 15 14 14 22 22 19 19 10
22 13 22 15 23 16 23 19 22 5 14 22 5 10 20 20 22 19 5 20 6 15 10 1
```

```
#Desgloso las acciones que se encuentran en cada neurona del mapa
```

```
1 which(map$unit.classif == 1)
```

```
[1] 78 79 157 278 289 303 316 319 349 376 424 501
```

```
1 which(map$unit.classif == 2)
```

```
[1] 8 11 18 19
```

```
1 which(map$unit.classif == 3)
```

```
[1] 1 2 3 4 5
```

```
1 which(map$unit.classif == 4)
```

```
[1] 20 28 31 35 36 40 42 46 76 365 388
```

```
1 which(map$unit.classif == 5)
```

```
[1] 24 41 55 60 84 105 141 203 224 245 262 271 276 286 326 429 445 487 490 496
```

```
1 which(map$unit.classif == 6)
```

```
[1] 6 72 95 130 287 385 498
```

```
1 which(map$unit.classif == 7)
```

```
[1] 49 53 80 85 117 119 138 145 148 149 150 159 160 166 168 210 211 225 235 238 280 344
377 421
```

```
1 which(map$unit.classif == 8)
```

```
[1] 22 23 26 30 32 34 38 39 44 45 51 54 56 57
```

```
1 which(map$unit.classif == 9)
```

```
[1] 7 9 10 13 14 15 16 17
```

```
1 which(map$unit.classif == 10)
```

```
[1] 37 97 110 120 181 193 194 196 199 206 209 277 295 331 334 339 342 359 380 386 406
409 431 444 477 491 500
```

```

1 which(map$unit.classif == 11)

[1] 12 231
1 which(map$unit.classif == 12)

[1] 47 71 106 108 137 162 169 184 222 242 251 255 284 285 313 327 378 402 412 443 468
1 which(map$unit.classif == 13)

[1] 126 172 299 306 345 348 479
1 which(map$unit.classif == 14)

[1] 29 48 50 58 68 82 90 101 102 124 147 155 174 183 185 191 214 215 253 254 268 269
296 307 318 330 362 368 391 400 419 432 441 459 467 471 472 488
1 which(map$unit.classif == 15)

[1] 66 67 87 93 103 109 135 151 161 208 252 258 270 272 275 329 336 338 350 351 364 375
383 399 405 411 418 427 430 438 442 461 465 470 481 499
1 which(map$unit.classif == 16)

[1] 70 83 113 134 153 164 170 173 177 197 198 219 223 236 244 250 281 288 305 320 328
340 374 389 426 483
1 which(map$unit.classif == 17)

[1] 21 25 116 156 221 273 300 315 325
1 which(map$unit.classif == 18)

[1] 61 62 63 64 89 91 92 100 104 112 114 118 125 128 132 136 178 179 186 188 201 212
228 234 246 256 291 301 312 323 333 335 356 361 382 403 422 433 454 455
1 which(map$unit.classif == 19)

[1] 43 65 77 88 107 122 131 133 140 192 202 243 248 249 257 264 266 274 353 354 366 392
423 428 440 452 457 462 475 476 485 495
1 which(map$unit.classif == 20)

[1] 73 121 142 176 189 216 217 230 232 260 265 298 302 322 371 372 407 410 415 439 450
451 464 492 493 497
1 which(map$unit.classif == 21)

[1] 59 96 123 127 129 139 158 165 175 182 195 218 226 261 293 314 324 332 343 346 357
369 370 381 387 408 413 425 435

```

```
1 which(map$unit.classif == 22)
```

```
[1] 74 75 94 98 205 240 241 247 259 263 282 283 304 310 311 317 341 358 360 363 373 384  
393 394 397 398 401 416 417 446 447 453 456 458 463 466 469 473 474 478 480 486 489  
494
```

```
1 which(map$unit.classif == 23)
```

```
[1] 33 69 81 99 111 143 144 146 152 154 167 171 187 190 200 204 207 213 227 233 237 239  
267 279 290 292 294 297 308 321 337 347 352 355 367 379 390 395 414 420 434 436 437  
448 460 482 484
```

```
1 which(map$unit.classif == 24)
```

```
[1] 27 52 115 163 180 220 396
```

```
1 which(map$unit.classif == 25)
```

```
[1] 86 229 309 404 449
```