

## Aviso al lector

Este post nace tras leer el artículo de Christian Oliva y Silviu Gabriel Tinjala, "Modeling Market States with Clustering and State Machines". Lo aquí desarrollado parte de una lectura atenta del mismo que recomiendo a quien haya tenido a bien acabar en esta web.

Mi intención no es menospreciar el trabajo de los autores, que considero inspirador, sino continuar con su línea de investigación y reforzar mi conocimiento en la materia.

Este y los artículos que seguirán se enmarcan en la investigación para mi Trabajo Final de Máster. Explicaré algunas ideas y técnicas empleadas, tratando de dar respuesta a esos "y si..." que interrumpen la lectura de un artículo interesante.

Nada de lo presente en esta web es consejo de inversión ni tiene por qué estar libre de errores.

---

## 1. Introducción

Este blog pretende ser útil tanto para especialistas como para no especialistas. Antes de entrar en materia, voy a contextualizar algunos términos.

### ¿Qué es un régimen de mercado?

Un régimen de mercado es un periodo extendido donde un activo muestra características homogéneas de retorno, volatilidad y comportamiento. No hablamos del estado de la economía en general, sino del comportamiento específico de un activo cotizado: un índice como el S&P 500, una acción individual, o un ETF sectorial.

Es cierto que el estado de un índice amplio como el S&P 500 contiene información sobre la economía, pero el objeto de este blog no es estudiar ciclos económicos sino el comportamiento de activos en bolsa.

### ¿Por qué importa identificar regímenes?

La pregunta clave es esta: Si cada régimen tiene unas determinadas características, que podemos conocer y estudiar en base a datos históricos, ¿no podríamos determinar en qué grado las circunstancias actuales se asemejan a uno u otro régimen? Y si conocemos que existen determinados regímenes diferentes, que los retornos que muestran cada uno de esos regímenes a futuro son diferentes... ¿Por qué no intentamos hacer una estrategia que incorpore el conocimiento del momento del ciclo estamos y nos libre de caídas o volatilidad indeseada?

### El enfoque clásico: All Weather de Ray Dalio

Una forma de abordar este problema es la diversificación estática. El famoso All Weather Portfolio de Bridgewater parte de la idea de que no sabemos qué régimen

vendrá, así que construimos una cartera que funcione razonablemente en todos ellos. Se diversifica entre activos que responden de forma diferente a distintos entornos: acciones para crecimiento, bonos para deflación, oro y commodities para inflación.

Este enfoque es todoterreno y defensivo. Su principio es el *Risk Parity*: asignar la misma cantidad de riesgo a cada tipo de activo. Un ejemplo simple: si la volatilidad del oro es 2 y de las acciones 4, dedico el doble de capital al oro que a las acciones para igualar la contribución al riesgo de la cartera de cada uno.

La limitación de este enfoque es que trata todos los momentos por igual. Si pudiéramos anticipar el régimen, podríamos sobreponer los activos que mejor funcionan en ese entorno específico.

### Nuestro enfoque: identificar el régimen y actuar en consecuencia

Intentamos extraer características del precio que, por sí mismas y sin depender de publicaciones macroeconómicas, nos permitan distinguir las fases de un activo. Podríamos comprar activos en régimen favorable y vender aquellos en situación contraria.

La idea es:

1. Identificar qué regímenes existen (usando datos históricos)
2. Determinar en qué régimen estamos (o hacia cuál vamos)
3. Ajustar la cartera según esa información

El primer paso — identificar regímenes — es un problema de agrupación. Queremos encontrar "clusters" de días con características similares. Para eso usamos técnicas de clustering, que es donde entra el artículo de Oliva y Tinjala.

### ¿Qué es el clustering?

Existen diferentes tipos de clustering que no vamos a detallar, pero la idea básica es la siguiente: son algoritmos que buscan agrupar puntos de forma que los de un mismo grupo sean similares entre sí y diferentes a los de otros grupos.

En este post usamos K-Means, uno de los algoritmos más conocidos. La mecánica es sencilla:

1. **Decides cuántos grupos (K) quieres.** Existen técnicas para elegir este número (método del codo, silhouette score) que no detallaré aquí.
2. **El algoritmo coloca K "centroides" en el espacio de características.** Un centroide es el punto central representativo de cada grupo.
3. **Cada observación se asigna al centroide más cercano.** La cercanía se mide típicamente con distancia euclíadiana.
4. **Los centroides se recalculan** como el promedio de los puntos asignados a ellos.
5. **Se repiten los pasos 3 y 4** hasta que las asignaciones no cambien.

El resultado: cada día queda asignado a un cluster, y cada cluster tiene un centroide que representa sus características típicas.

Para entenderlo visualmente recomiendo [este vídeo](#). Para una lectura más detallada, [este artículo](#)

### **Aplicado a nuestro problema:**

Cada día de mercado tiene un vector de características derivadas del precio (volatilidad, retornos, etc.). K-Means agrupa los días que tienen características similares. Si elegimos K=4, por ejemplo, podríamos obtener clusters que interpretamos como "expansión", "contracción", "recuperación" y "crisis" — aunque los nombres se los ponemos nosotros después de ver qué características tiene cada centroide.

## **2. Implementación**

Bien, ya he explicado lo que es un régimen de mercado y cómo los algoritmos de clustering nos permiten capturar puntos de tiempo con comportamientos diferentes.

Pongámonos manos a la obra.

Lo primero que tenemos que hacer es descargarnos datos históricos de precios del índice SP500 y obtener características del movimiento del mismo.

La primera pregunta es:

**¿Qué características del precio me ayudan a determinar el estado del mercado?**