

Construcción de Carteras usando Machine Learning

Fernando de la Calle Silos

21 de julio de 2020

BME Inntech
Instituto BME

1. Introducción
2. Consideraciones Previas
3. Modern Portfolio Theory
4. Introducción al Machine Learning
5. Clustering Jerárquico Aglomerativo
6. Hierarchical Risk Parity
7. Conclusiones

Introducción



- Máster en Inteligencia Artificial aplicada a los Mercados Financieros (mIA-X).



- Máster en Inteligencia Artificial aplicada a los Mercados Financieros (mIA-X).
- https://www.institutobme.es/esp/curso/M%C3%A1ster_en_Inteligencia_Artificial_aplicada_a_los_Mercados_Financieros__mIA_X_.aspx



- El laboratorio nace en Sep-2018 como un equipo multidisciplinar de alto rendimiento. Sus 5 integrantes cuentan con perfiles diversos: Matemático, PhD Computer Science, PhD Multimedia and Communications, Licenciado en Derecho y Dirección de empresas.



- El laboratorio nace en Sep-2018 como un equipo multidisciplinar de alto rendimiento. Sus 5 integrantes cuentan con perfiles diversos: Matemático, PhD Computer Science, PhD Multimedia and Communications, Licenciado en Derecho y Dirección de empresas.
- El objetivo del laboratorio es aplicar la potencia de la Inteligencia Artificial a los Mercados Financieros, buscando dar un salto cualitativo en el desarrollo de aplicaciones bursátiles.



- El laboratorio nace en Sep-2018 como un equipo multidisciplinar de alto rendimiento. Sus 5 integrantes cuentan con perfiles diversos: Matemático, PhD Computer Science, PhD Multimedia and Communications, Licenciado en Derecho y Dirección de empresas.
- El objetivo del laboratorio es aplicar la potencia de la Inteligencia Artificial a los Mercados Financieros, buscando dar un salto cualitativo en el desarrollo de aplicaciones bursátiles.
- <https://www.bmeinntech.es/esp/Home>

Consideraciones Previas

Consideraciones Previas

- Algunas consideraciones antes de empezar:

- Algunas consideraciones antes de empezar:
 - Sesgo de supervivencia

Consideraciones Previas

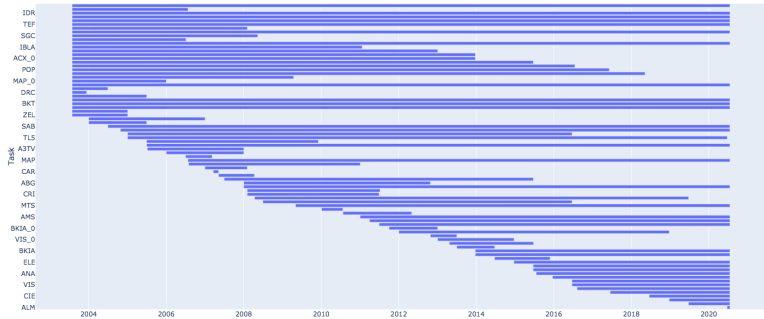
- Algunas consideraciones antes de empezar:
 - Sesgo de supervivencia
 - Cómo saber si nuestra teoría es correcta

Consideraciones Previas

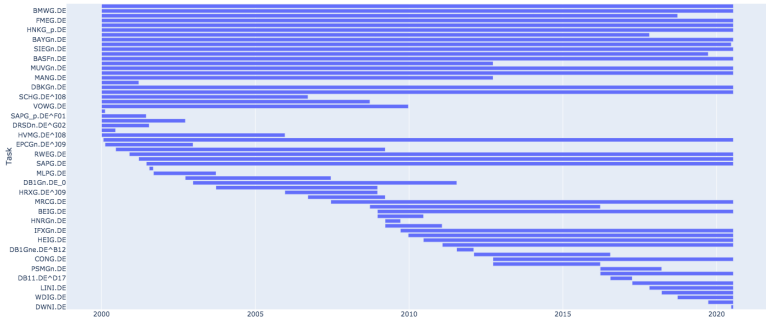
- Algunas consideraciones antes de empezar:
 - Sesgo de supervivencia
 - Cómo saber si nuestra teoría es correcta
 - Sesgo de selección

Sesgo de Supervivencias

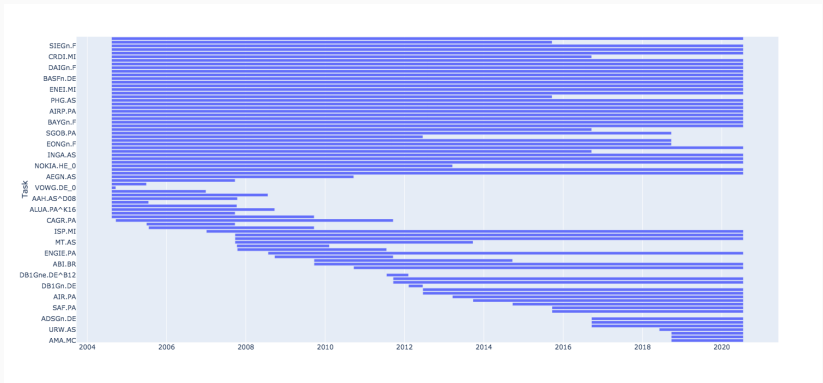
- Las componentes del índice cambian con el tiempo.



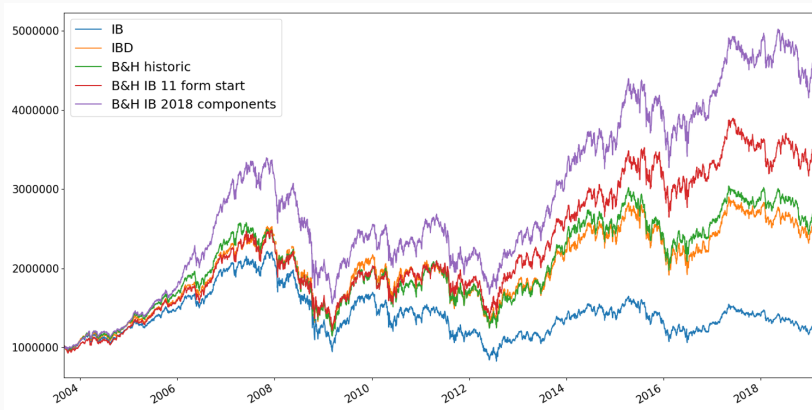
Sesgo de Supervivencia: DAX



Sesgo de Supervivencia: EUROSTOXX 50



Sesgo de Supervivencia: IBEX



Cómo saber si nuestra teoría es correcta.

- Algunas métricas que podemos observar pueden ser:

Cómo saber si nuestra teoría es correcta.

- Algunas métricas que podemos observar pueden ser:
 - Rendimiento compuesto

Cómo saber si nuestra teoría es correcta.

- Algunas métricas que podemos observar pueden ser:
 - Rendimiento compuesto
 - Rendimiento por año

Cómo saber si nuestra teoría es correcta.

- Algunas métricas que podemos observar pueden ser:
 - Rendimiento compuesto
 - Rendimiento por año
 - Drawdown

Cómo saber si nuestra teoría es correcta.

- Algunas métricas que podemos observar pueden ser:
 - Rendimiento compuesto
 - Rendimiento por año
 - Drawdown
 - Rolling Return

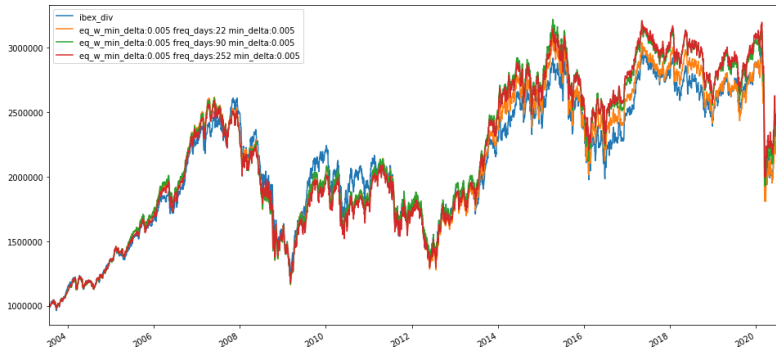
Cómo saber si nuestra teoría es correcta.

- Algunas métricas que podemos observar pueden ser:
 - Rendimiento compuesto
 - Rendimiento por año
 - Drawdown
 - Rolling Return
 - Rolling Return Distribution

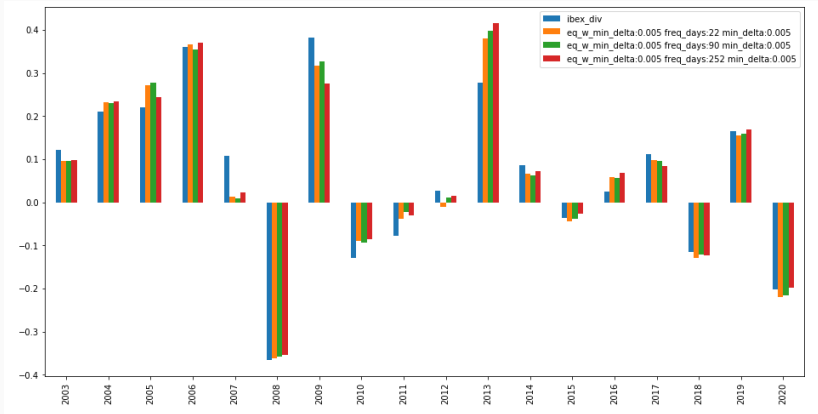
Cómo saber si nuestra teoría es correcta.

- Algunas métricas que podemos observar pueden ser:
 - Rendimiento compuesto
 - Rendimiento por año
 - Drawdown
 - Rolling Return
 - Rolling Return Distribution
 - Sharpe ratio, IR, tracking error, ulcer index...

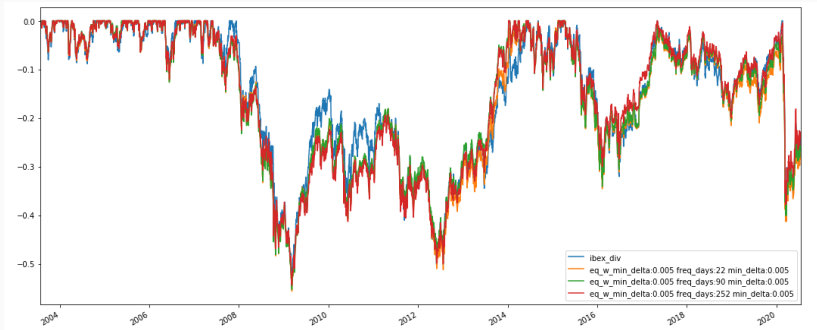
Rendimiento compuesto



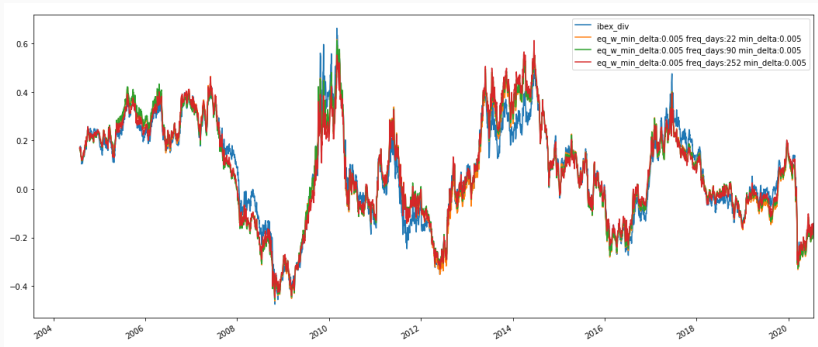
Rendimiento por año



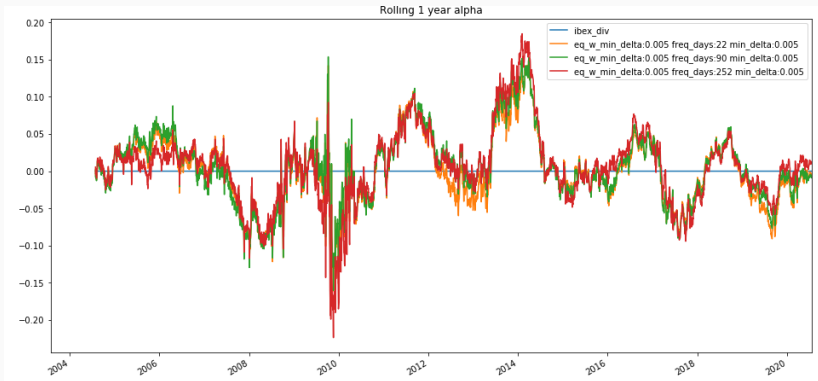
Drawdown



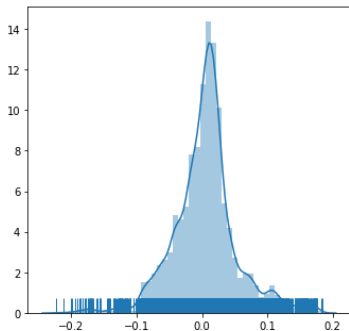
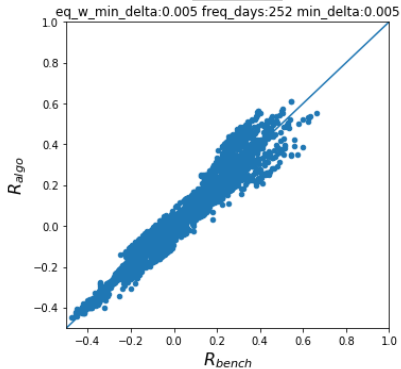
Rolling Return 1 year



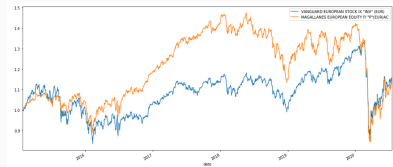
Rolling Alpha



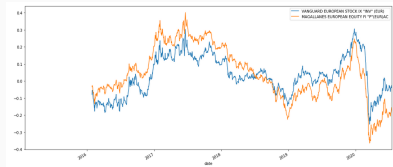
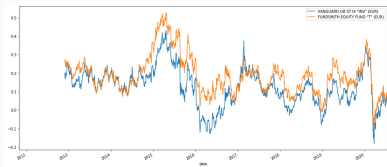
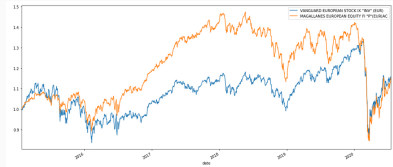
Rolling Return Scatter



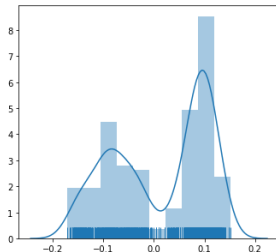
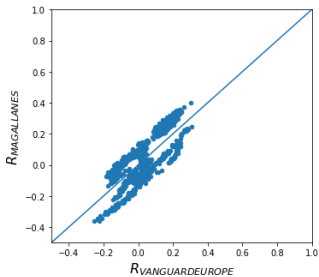
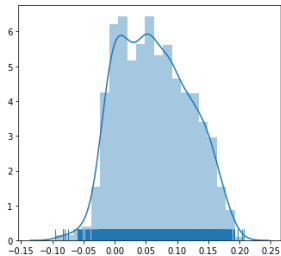
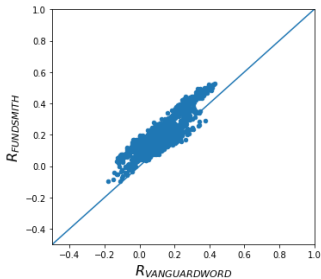
Ejemplo con fondos



Ejemplo con fondos



Ejemplo con fondos



- Sobreajustar al pasado no es la solución.

- Sobreajustar al pasado no es la solución.
- Cuando tenemos parámetros o modelos necesitamos una forma de entrenar los mismos.

Sesgo de Selección

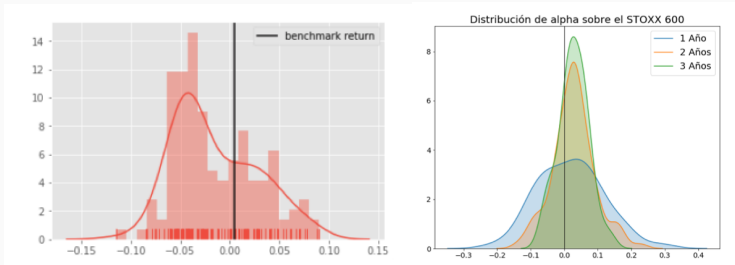
- Sobreajustar al pasado no es la solución.
- Cuando tenemos parámetros o modelos necesitamos una forma de entrenar los mismos.
- Bootstrapping, validación combinatoria, datos sintéticos, etc [1, 2].

Sesgo de Selección

- Sobreajustar al pasado no es la solución.
- Cuando tenemos parámetros o modelos necesitamos una forma de entrenar los mismos.
- Bootstrapping, validación combinatoria, datos sintéticos, etc [1, 2].

Sesgo de Selección

- Sobreajustar al pasado no es la solución.
- Cuando tenemos parámetros o modelos necesitamos una forma de entrenar los mismos.
- Bootstrapping, validación combinatoria, datos sintéticos, etc [1, 2].



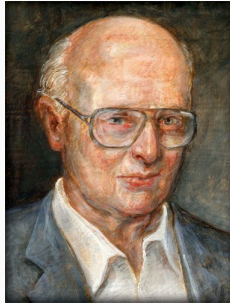
Modern Portfolio Theory

Modern Portfolio Theory

- Modern portfolio theory (MPT) o mean-variance analysis.

Modern Portfolio Theory

- Modern portfolio theory (MPT) o mean-variance analysis.
- Propuesto por Harry Markowitz en 1952 [3].



Novel 1990

Modern Portfolio Theory

- Modern portfolio theory (MPT) o mean-variance analysis.
- Propuesto por Harry Markowitz en 1952 [3].

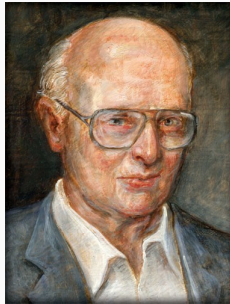


Novel 1990

- Obtener una cartera de activos que maximizan el retorno para un nivel dado de riesgo.

Modern Portfolio Theory

- Modern portfolio theory (MPT) o mean-variance analysis.
- Propuesto por Harry Markowitz en 1952 [3].



Novel 1990

- Obtener una cartera de activos que maximizan el retorno para un nivel dado de riesgo.
- Basado en el retorno medio esperado y desviación típica esperada.

- Dada una serie de retornos para cada activo:

- Dada una serie de retornos para cada activo:
- Encontrar un vector de pesos \mathbf{w} de dimensiones $n \times 1$, donde n es el número de activos. Los pesos tienen que cumplir que $w_i \geq 0$, $i = 1, \dots, n$ y $\sum_{i=1}^n w_i = 1$ dado que solo permitimos estar en largo.

Modern Portfolio Theory

- Dada una serie de retornos para cada activo:
- Encontrar un vector de pesos \mathbf{w} de dimensiones $n \times 1$, donde n es el número de activos. Los pesos tienen que cumplir que $w_i \geq 0$, $i = 1, \dots, n$ y $\sum_{i=1}^n w_i = 1$ dado que solo permitimos estar en largo.
- El retorno y la desviación típica de este portfolio se calcula como:

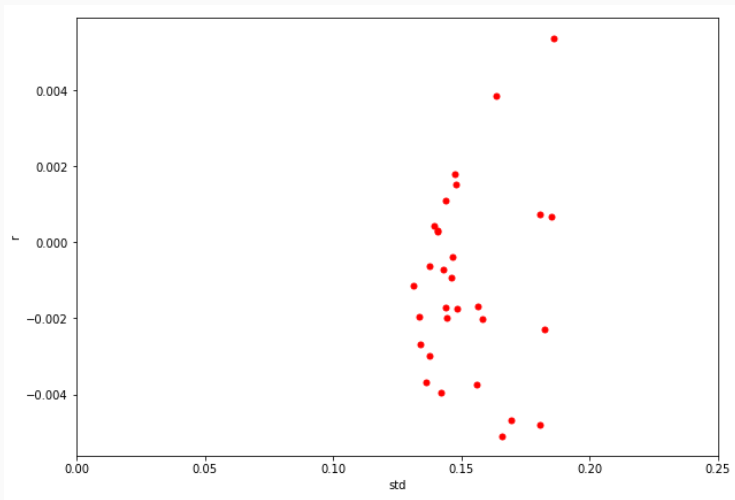
Modern Portfolio Theory

- Dada una serie de retornos para cada activo:
- Encontrar un vector de pesos \mathbf{w} de dimensiones $n \times 1$, donde n es el número de activos. Los pesos tienen que cumplir que $w_i \geq 0$, $i = 1, \dots, n$ y $\sum_{i=1}^n w_i = 1$ dado que solo permitimos estar en largo.
- El retorno y la desviación típica de este portfolio se calcula como:
 - $R_p = \mathbf{w}^T \mu$ donde μ es un vector con la media de los retornos diarios de los activos, con dimensiones $n \times 1$.

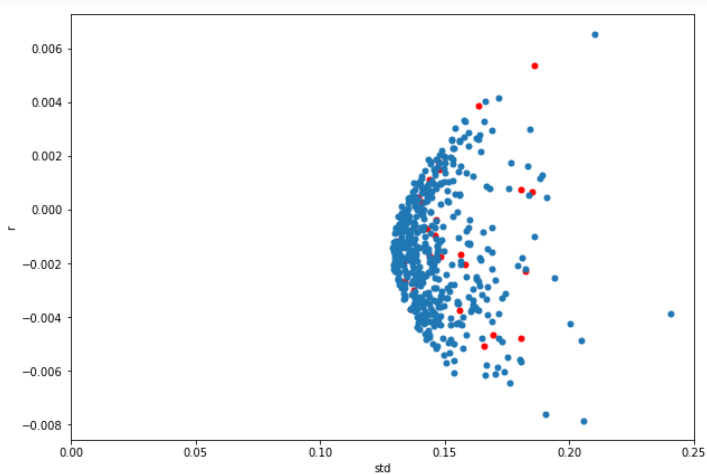
Modern Portfolio Theory

- Dada una serie de retornos para cada activo:
- Encontrar un vector de pesos \mathbf{w} de dimensiones $n \times 1$, donde n es el número de activos. Los pesos tienen que cumplir que $w_i \geq 0$, $i = 1, \dots, n$ y $\sum_{i=1}^n w_i = 1$ dado que solo permitimos estar en largo.
- El retorno y la desviación típica de este portfolio se calcula como:
 - $R_p = \mathbf{w}^T \mu$ donde μ es un vector con la media de los retornos diarios de los activos, con dimensiones $n \times 1$.
 - $\sigma_p = \sqrt{\mathbf{w}^T \Sigma \mathbf{w}}$, donde Σ es la matriz de covarianzas de los retornos, con dimensiones $n \times n$

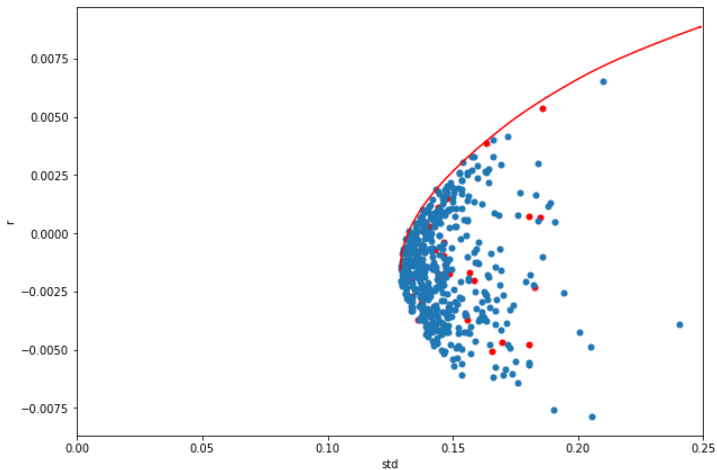
Modern Portfolio Theory



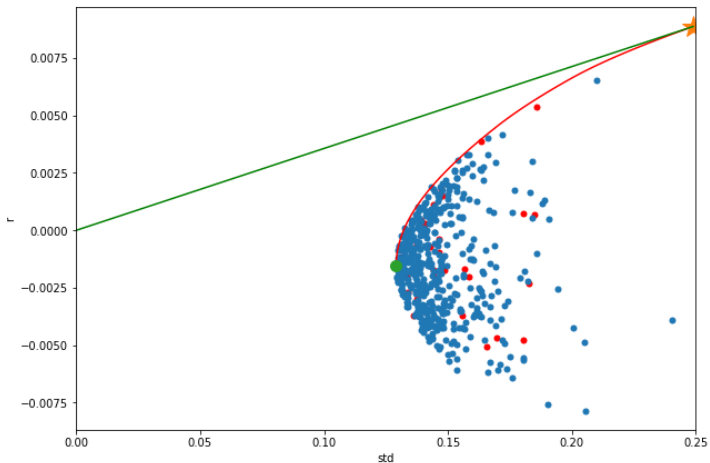
Modern Portfolio Theory



Modern Portfolio Theory



Modern Portfolio Theory



Frontera eficiente

- El problema de la frontera eficiente se puede tratar como un problema de optimización cuadrática.

$$\text{maximize} \quad \mathbf{w}^T \boldsymbol{\mu} - \gamma \mathbf{w}^T \boldsymbol{\Sigma} \mathbf{w}$$

$$\text{subject to} \quad w_i \geq 0, \quad i = 1, \dots, n$$

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

Frontera eficiente

- El problema de la frontera eficiente se puede tratar como un problema de optimización cuadrática.

$$\text{maximize} \quad \mathbf{w}^T \boldsymbol{\mu} - \gamma \mathbf{w}^T \boldsymbol{\Sigma} \mathbf{w}$$

$$\text{subject to} \quad w_i \geq 0, \quad i = 1, \dots, n$$

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

- Donde \mathbf{w} es el vector de pesos a optimizar, $\mathbf{w}^T \boldsymbol{\mu}$ es el retorno del portfolio, $\mathbf{w}^T \boldsymbol{\Sigma} \mathbf{w}$ es la varianza del portfolio y γ es un parámetro de aversión al riesgo.

Frontera eficiente

- El problema de la frontera eficiente se puede tratar como un problema de optimización cuadrática.

$$\text{maximize} \quad \mathbf{w}^T \boldsymbol{\mu} - \gamma \mathbf{w}^T \boldsymbol{\Sigma} \mathbf{w}$$

$$\text{subject to} \quad w_i \geq 0, \quad i = 1, \dots, n$$

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

- Donde \mathbf{w} es el vector de pesos a optimizar, $\mathbf{w}^T \boldsymbol{\mu}$ es el retorno del portfolio, $\mathbf{w}^T \boldsymbol{\Sigma} \mathbf{w}$ es la varianza del portfolio y γ es un parámetro de aversión al riesgo.
- Variando γ obtenemos los diferentes puntos de la frontera. Si $\gamma = 0$ maximizamos solo el retorno y si $\gamma \rightarrow \infty$ obtenemos el portfolio de mínimo riesgo. Variando γ obtenemos el *risk-return-trade-off*.

Critical Line Algorithm (CLA)

- Harry Markowitz desarrolló en 1956 el Critical Line Algorithm (CLA) [4].

Critical Line Algorithm (CLA)

- Harry Markowitz desarrolló en 1956 el Critical Line Algorithm (CLA) [4].
- CLA es un proceso de optimización cuadrática que encuentra la solución exacta después de un número de pasos conocido.

Critical Line Algorithm (CLA)

- Harry Markowitz desarrolló en 1956 el Critical Line Algorithm (CLA) [4].
- CLA es un proceso de optimización cuadrática que encuentra la solución exacta después de un número de pasos conocido.
- Solventa algunos problemas de los optimizadores cuadráticos al aplicarse a este problema.

Critical Line Algorithm (CLA)

- Harry Markowitz desarrolló en 1956 el Critical Line Algorithm (CLA) [4].
- CLA es un proceso de optimización cuadrática que encuentra la solución exacta después de un número de pasos conocido.
- Solventa algunos problemas de los optimizadores cuadráticos al aplicarse a este problema.
- Se puede encontrar un desarrollo open source en [5].

DEMO

- Da lugar a carteras óptimas en el periodo usado, *in sample*.

- Da lugar a carteras óptimas en el periodo usado, *in sample*.
- Tiende a obtener un rendimiento malo *out-of-sample* (en algunos casos peor que eqw).

Problemas

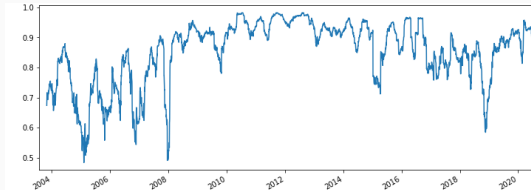
- Los retornos no se pueden estimar con precisión.

Problemas

- Los retornos no se pueden estimar con precisión.
- Requiere la inversión de la matriz de covarianzas (muy cercana a ser singular) [2].

Problemas

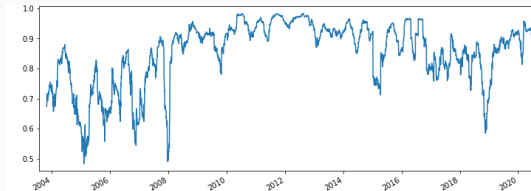
- Los retornos no se pueden estimar con precisión.
- Requiere la inversión de la matriz de covarianzas (muy cercana a ser singular) [2].
- Las correlaciones de los activos son muy ruidosas y cambiantes.



3 meses roll corr BBVA-SAN

Problemas

- Los retornos no se pueden estimar con precisión.
- Requiere la inversión de la matriz de covarianzas (muy cercana a ser singular) [2].
- Las correlaciones de los activos son muy ruidosas y cambiantes.

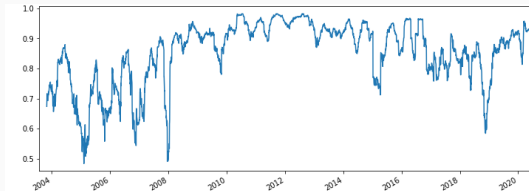


3 meses roll corr BBVA-SAN

- Un pequeño cambio en la matriz de correlación da lugar a un gran alteración de la cartera generada.

Problemas

- Los retornos no se pueden estimar con precisión.
- Requiere la inversión de la matriz de covarianzas (muy cercana a ser singular) [2].
- Las correlaciones de los activos son muy ruidosas y cambiantes.



3 meses roll corr BBVA-SAN

- Un pequeño cambio en la matriz de correlación da lugar a un gran alteración de la cartera generada.
- Riesgo idiosincrásico: en muchas ocasiones produce carteras muy concentradas.

Resultados



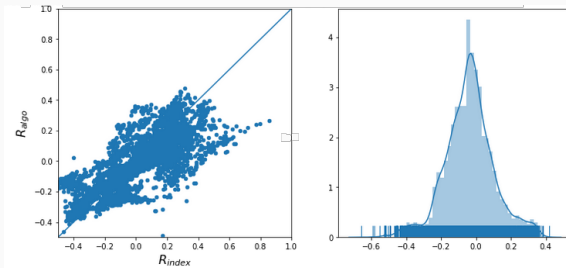
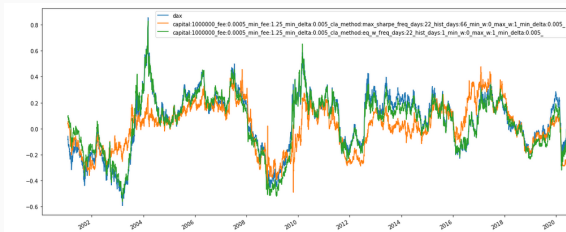
Resultados



Resultados



Resultados



Introducción al Machine Learning

- Aprendizaje Supervisado: Los ejemplos vienen con una variable de salida o etiqueta que queremos aprender.

Machine Learning

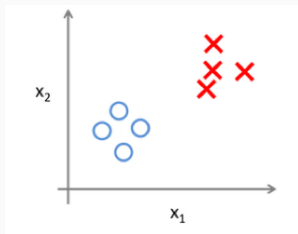
- Aprendizaje Supervisado: Los ejemplos vienen con una variable de salida o etiqueta que queremos aprender.
- Aprendizaje No Supervisado: No existe una variable o concepto objetivo. Nos interesa descubrir patrones o relaciones que expliquen los datos.

Machine Learning

- Aprendizaje Supervisado: Los ejemplos vienen con una variable de salida o etiqueta que queremos aprender.
- Aprendizaje No Supervisado: No existe una variable o concepto objetivo. Nos interesa descubrir patrones o relaciones que expliquen los datos.

Machine Learning

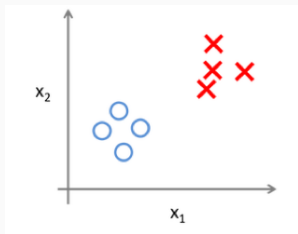
- Aprendizaje Supervisado: Los ejemplos vienen con una variable de salida o etiqueta que queremos aprender.
- Aprendizaje No Supervisado: No existe una variable o concepto objetivo. Nos interesa descubrir patrones o relaciones que expliquen los datos.



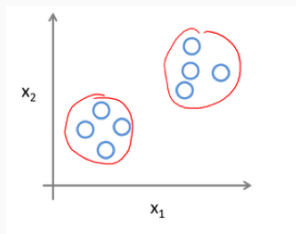
Aprendizaje Supervisado

Machine Learning

- Aprendizaje Supervisado: Los ejemplos vienen con una variable de salida o etiqueta que queremos aprender.
- Aprendizaje No Supervisado: No existe una variable o concepto objetivo. Nos interesa descubrir patrones o relaciones que expliquen los datos.



Aprendizaje Supervisado



Aprendizaje No Supervisado

Clustering Jerárquico Aglomerativo

- Agrupamiento jerárquico es un método de análisis de grupos puntuales, el cual busca construir una jerarquía de grupos

Clustering Jerárquico Aglomerativo

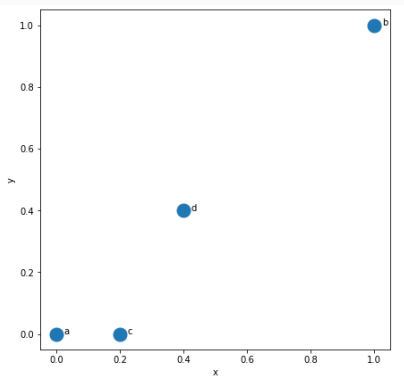
- Agrupamiento jerárquico es un método de análisis de grupos puntuales, el cual busca construir una jerarquía de grupos
- En el aglomerativo se sigue un acercamiento ascendente; cada observación comienza en su propio grupo, y los pares de grupos son mezclados mientras uno sube en la jerarquía.

Clustering Jerárquico Aglomerativo

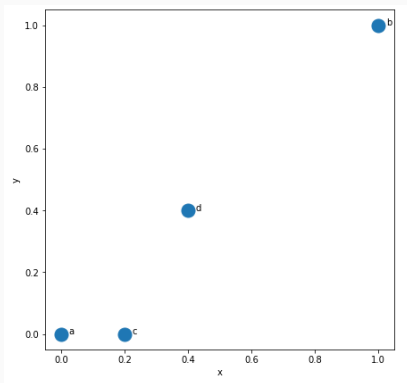
	x	y
a	0.0	0.0
b	1.0	1.0
c	0.2	0.0
d	0.4	0.4

Clustering Jerárquico Aglomerativo

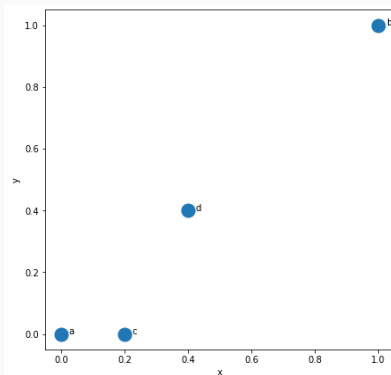
	x	y
a	0.0	0.0
b	1.0	1.0
c	0.2	0.0
d	0.4	0.4



Clustering Jerárquico Aglomerativo

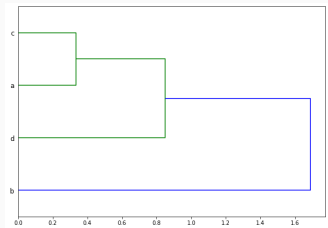
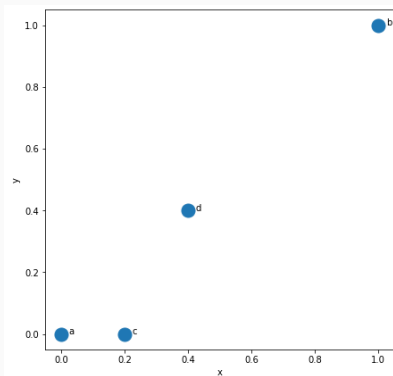


Clustering Jerárquico Aglomerativo



	a	b	c	d
a	0.000000	1.414214	0.200000	0.565685
b	1.414214	0.000000	1.280625	0.848528
c	0.200000	1.280625	0.000000	0.447214
d	0.565685	0.848528	0.447214	0.000000

Clustering Jerárquico Aglomerativo



Hierarchical Risk Parity

- Propuesto por Marcos Lopez de Prado en [1, 6].

Hierarchical Risk Parity

- Propuesto por Marcos Lopez de Prado en [1, 6].
- Es un método de construcción de carteras que no necesita invertir la matriz de covarianza .

Hierarchical Risk Parity

- Propuesto por Marcos Lopez de Prado en [1, 6].
- Es un método de construcción de carteras que no necesita invertir la matriz de covarianza .
- Obtiene la jerarquía mediante clustering jerárquico aglomerativo y diversifica sobre los distintos clusters de forma igualitaria en cuanto al riesgo.

1. Trasforma la matriz de correlaciones en una medida de distancia:

$$d_{i,j} = \sqrt{\frac{1}{2}(1 - \rho_{i,j})}$$

Hierarchical Risk Parity

1. Transforma la matriz de correlaciones en una medida de distancia:

$$d_{i,j} = \sqrt{\frac{1}{2}(1 - \rho_{i,j})}$$

2. Aplicar el algoritmo de clustering jerárquico sobre esta matriz

Hierarchical Risk Parity

1. Transforma la matriz de correlaciones en una medida de distancia:

$$d_{i,j} = \sqrt{\frac{1}{2}(1 - \rho_{i,j})}$$

2. Aplicar el algoritmo de clustering jerárquico sobre esta matriz
3. Quasi-Diagonalization: reorganizar la matriz de forma que los elementos correlados entre si queden juntos.

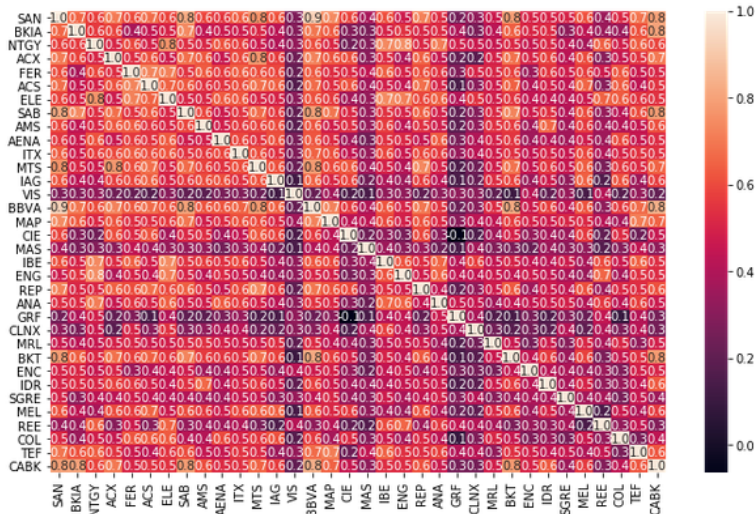
Hierarchical Risk Parity

1. Transforma la matriz de correlaciones en una medida de distancia:

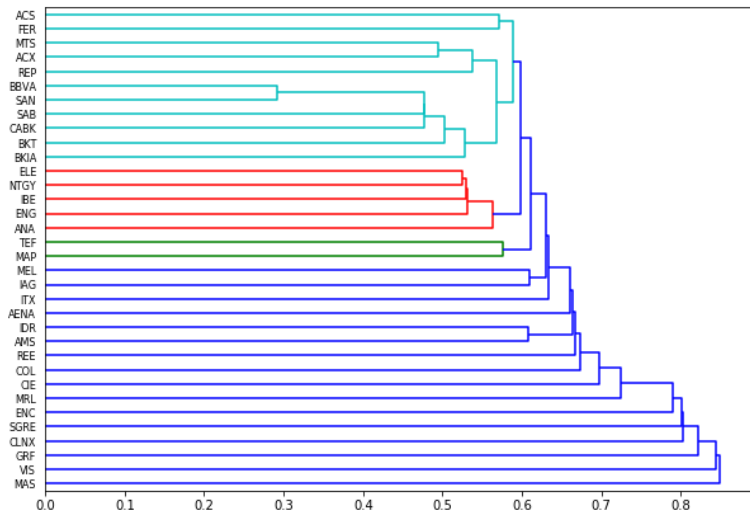
$$d_{i,j} = \sqrt{\frac{1}{2}(1 - \rho_{i,j})}$$

2. Aplicar el algoritmo de clustering jerárquico sobre esta matriz
3. Quasi-Diagonalization: reorganizar la matriz de forma que los elementos correlados entre si queden juntos.
4. Recursive Bisection: repartir top-down los pesos en función de la varianza de cada sub-cluster (risk parity). Solo los elemento de cada cluster compiten por los pesos.

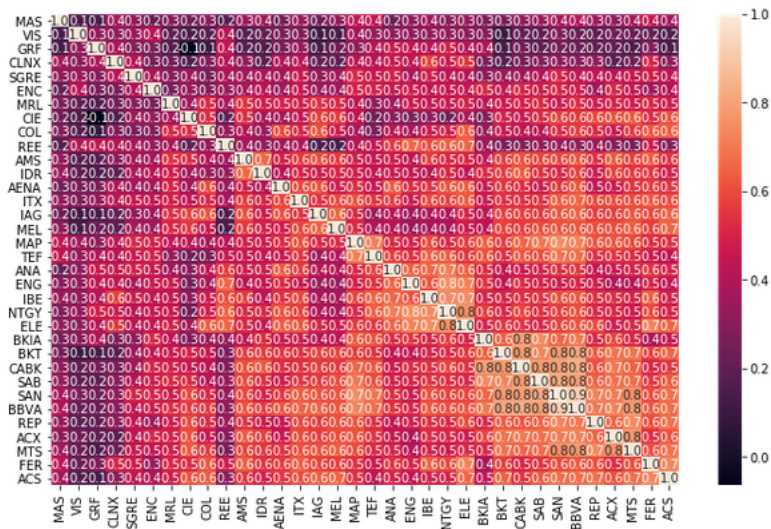
IBEX Corr



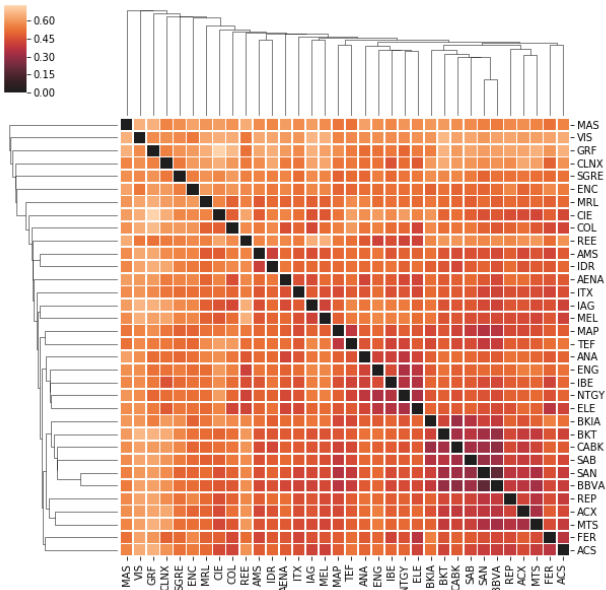
Dendrograma



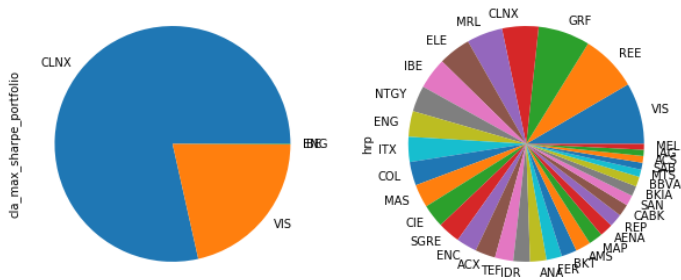
Quasi-Diagonalization IBEX Corr



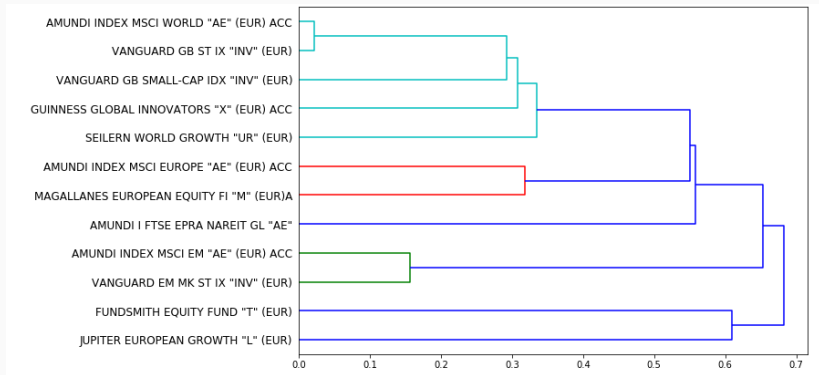
Quasi-Diagonalization IBEX Corr



Carteras generadas

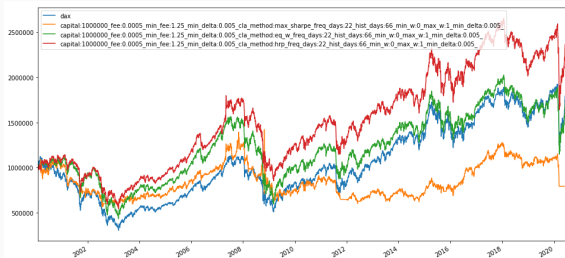


Para una cartera de fondos

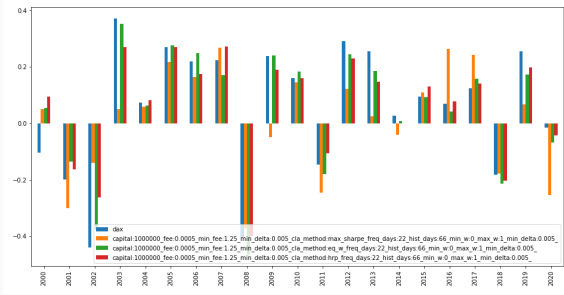


DEMO

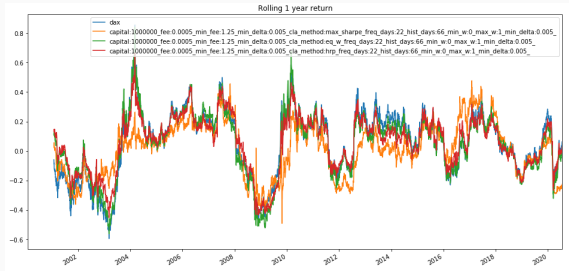
Resultados



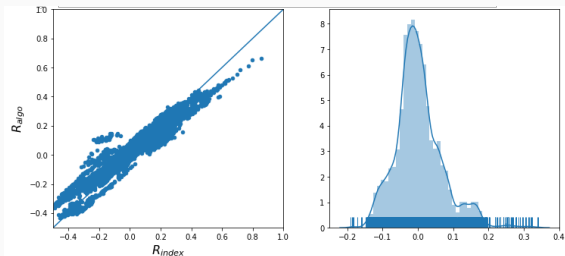
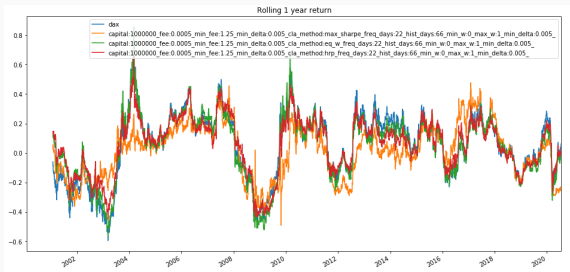
Resultados



Resultados



Resultados



Conclusiones

- Markowitz es la cartera óptima *in-sample*.

Conclusiones

- Markowitz es la cartera óptima *in-sample*.
- Como hemos visto tiene muchos inconvenientes asociados.

- Markowitz es la cartera óptima *in-sample*.
- Como hemos visto tiene muchos inconvenientes asociados.
- El uso de técnicas de machine learning nos puede valer para analizar, como para construir carteras.

Conclusiones

- Markowitz es la cartera óptima *in-sample*.
- Como hemos visto tiene muchos inconvenientes asociados.
- El uso de técnicas de machine learning nos puede valer para analizar, como para construir carteras.
- Hemos visto uno de los algoritmos de clustering más sencillos.

Conclusiones

- Markowitz es la cartera óptima *in-sample*.
- Como hemos visto tiene muchos inconvenientes asociados.
- El uso de técnicas de machine learning nos puede valer para analizar, como para construir carteras.
- Hemos visto uno de los algoritmos de clustering más sencillos.
- Su aplicación nos permite obtener carteras más diversificadas y que tienden a funcionar mejor *out-of-sample*.

¿Preguntas?

Referencias

- [1] M. L. de Prado, *Advances in Financial Machine Learning*. Wiley Publishing, 1st ed., 2018.
- [2] M. M. López de Prado, *Machine Learning for Asset Managers*. Elements in Quantitative Finance, Cambridge University Press, 2020.
- [3] H. Markowitz, "Portfolio selection," *Journal of Finance*, vol. 7, no. 1, pp. 77–91, 1952.
- [4] H. Markowitz, "The optimization of a quadratic function subject to linear constraints," *Naval Research Logistics Quarterly*, vol. 3, no. 1a2, pp. 111–133, 1956.

- [5] D. H. Bailey and M. López de Prado, “An open-source implementation of the critical-line algorithm for portfolio optimization,” 2013. Available at SSRN: <https://ssrn.com/abstract=2197616>.
- [6] M. López de Prado, “Building diversified portfolios that outperform out-of-sample,” *Journal of Portfolio Management*, 2016. Available at SSRN: <https://ssrn.com/abstract=2708678>.