Selección de Portafolios Óptimos Usando Clusters

DJ ES

21 de noviembre de 2013

Contenido

Motivación

Definimos los invariantes como:

$$X_{i,t,\tau} = ln(\frac{P_{i,t}}{P_{i,t-\tau}}) \tag{1}$$

Donde $P_{i,t}$ es el precio de una acción en el momento t en un horizonte de estimación de tamaño τ.

Motivación

El problema clásico de selección de portafolios se puede escribir como:

$$\min w^T \Omega w \tag{2}$$

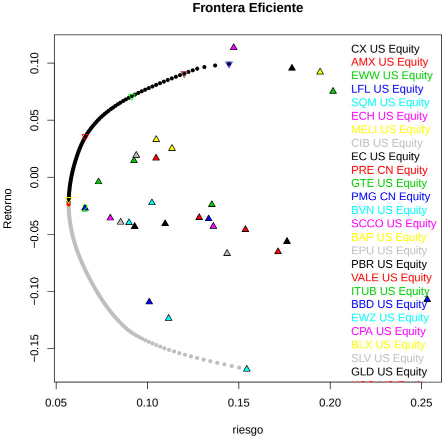
Sujeto a:

$$\begin{aligned} \sum_i \mu_i w_i &= \bar{r} \\ \sum_i w_i &= 1 \\ w_i^{min} \leq w_i &\leq w_i^{max} \end{aligned} \tag{3}$$

Donde w es un vector de ponderaciones, $\bar{r}$ es una constante, μ y Ω son la matriz de medias y las covarianzas de los invariantes X_i.

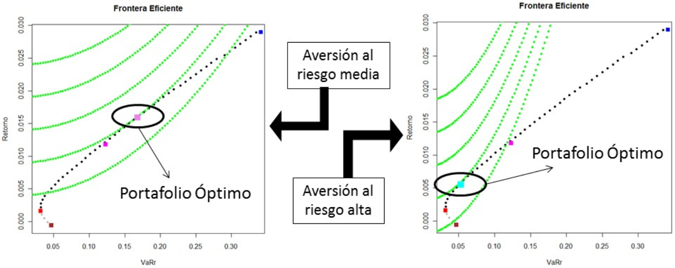
Motivación

Moviendo valores de $\bar{r}$ se pueden graficar en el espacio de retorno($\sum_i \mu_i w_i$) y varianza($w^T \Omega w$) las diferentes soluciones del problema descrito anteriormente, lo que se conoce como frontera eficiente:



Motivación

Normalmente se elige un portafolio basado en un perfil de aversión al riesgo (usando una función de utilidad sobre el espacio de retorno y varianza). Y se elige el portafolio que haga tangencia con la función de utilidad:



Motivación

Dado que μ y Ω no son conocidos, es necesario usar estimadores para estimarlos. Entre otros, algunos estimadores famosos son:

- Estimadores no-paramétricos.
- Estimadores de *shrinkage*.
- Black Litterman.

Motivación

Cuando se permiten ventas en corto $w_i^{min} < 0$ (Una venta en corto consiste en vender una acción sin tenerla, por lo tanto si el precio de la acción cae se obtiene ganancia dado que se puede comprar a un precio mas bajo) idealmente se espera que el portafolio tenga cierto cubrimiento de mercado.
El cubrimiento de mercado se refiere a que cuando empieza a caer el precio de las acciones en general, la caída de valor del portafolio en posiciones largas ($w_i > 0$) se vea compensado por aumentos en el valor del portafolio en las posiciones cortas ($w_i < 0$).

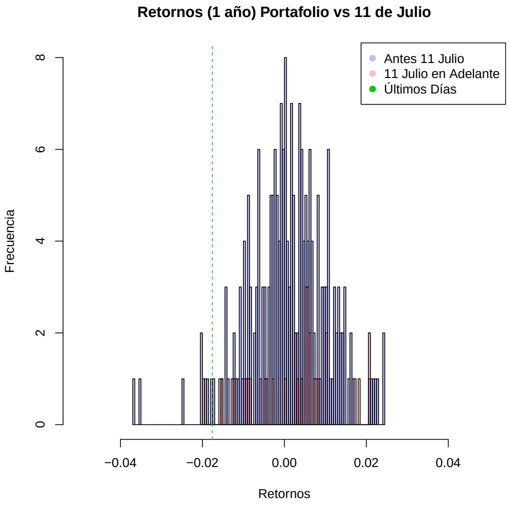
Motivación

En escenarios de crisis, la matriz de varianza-covarianza estimada presenta un comportamiento inestable. Por lo que en algunos casos los cubrimientos del portafolio no funcionaran en la dirección deseada, por ejemplo:

- Las correlaciones se invierten o se vuelven 1, y cuando las posiciones largas pierden valor las cortas también.

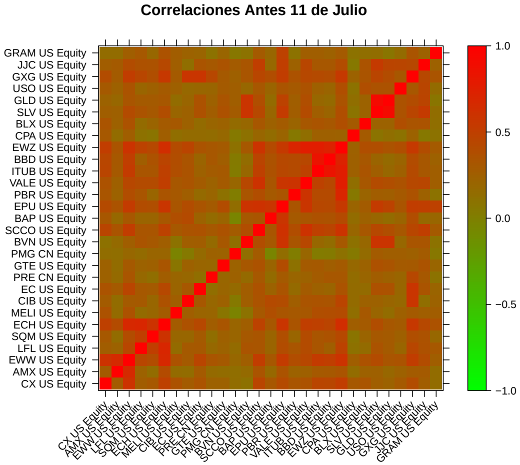
Motivación

Ejemplo distribución retorno diario portafolio durante una crisis:

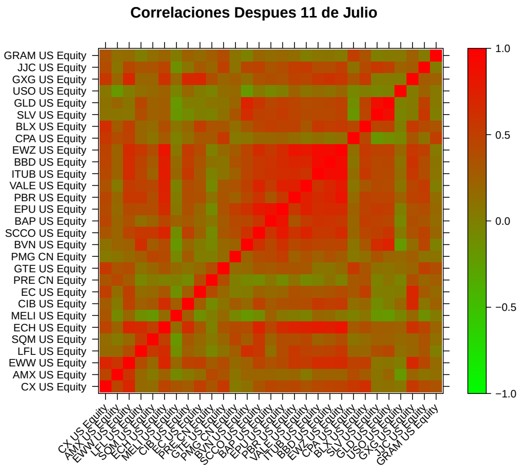


Se muestra el cambio de la matriz de varianza-covarianza durante el periodo de crisis.

Motivación

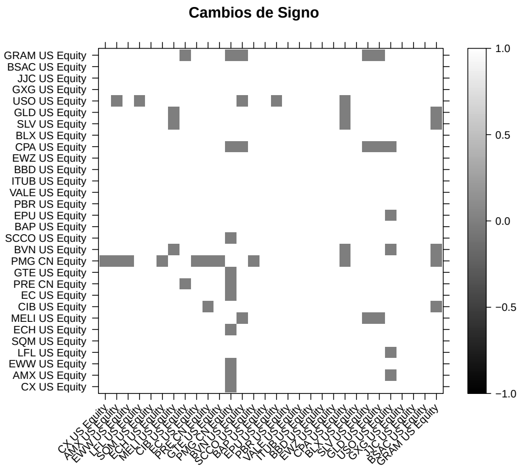

Correlaciones Antes 11 de Julio
GRAM US Equity
JJC US Equity
GXG US Equity
USO US Equity
GLD US Equity
SLV US Equity
BLX US Equity
CPA US Equity
EWZ US Equity
BBD US Equity
ITUB US Equity
VALE US Equity
PBR US Equity
EPU US Equity
BAP US Equity
SCCO US Equity
BVN US Equity
PMG CN Equity
GTE US Equity
PRE CN Equity
EC US Equity
CIB US Equity
MELI US Equity
ECH US Equity
SQM US Equity
LFL US Equity
EWW US Equity
AMX US Equity
CX US Equity
1.0
0.5
0.0
−0.5
−1.0

Motivación


Correlaciones Despues 11 de Julio
GRAM US Equity
JJC US Equity
GXG US Equity
USO US Equity
GLD US Equity
SLV US Equity
BLX US Equity
CPA US Equity
EWZ US Equity
BBD US Equity
ITUB US Equity
VALE US Equity
PBR US Equity
EPU US Equity
BAP US Equity
SCCO US Equity
BVN US Equity
PMG CN Equity
GTE US Equity
PRE CN Equity
EC US Equity
CIB US Equity
MELI US Equity
ECH US Equity
SQM US Equity
LFL US Equity
EWW US Equity
AMX US Equity
CX US Equity
1.0
0.5
0.0
−0.5
−1.0

El cambio es tan drástico que algunas covarianzas cambian de signo:

Motivación


Cambios de Signo
GRAM US Equity
BSAC US Equity
JJC US Equity
GXG US Equity
USO US Equity
GLD US Equity
SLV US Equity
BLX US Equity
CPA US Equity
EWZ US Equity
BBD US Equity
ITUB US Equity
VALE US Equity
PBR US Equity
EPU US Equity
BAP US Equity
SCCO US Equity
BVN US Equity
PMG CN Equity
GTE US Equity
PRE CN Equity
EC US Equity
CIB US Equity
MELI US Equity
ECH US Equity
SQM US Equity
LFL US Equity
EWW US Equity
AMX US Equity
CX US Equity
1.0
0.5
0.0
−0.5
−1.0

Dada la estructura de la estimación de la matriz de varianza-covarianza, el método de selección de portafolios clásico parece destinado a sufrir grandes perdidas en periodos de crisis.

Motivación

Para evitar esto proponemos agrupar en *clusters* los diferentes activos y luego proponer una estrategia de selección de portafolio, teniendo en mente los siguientes objetivos:

- Que los activos dentro de un cluster presenten un comportamiento *similar*.
- Que la metodología de estimación de los clusters sea relativamente suave.
- Establecer como objeto de inversión los clusters y no los activos en sí.

Contenido

Medidas

Para establecer una estrategia de clustering es necesario definir la similitud o distancia entre dos activos de tal forma que durante los escenarios de crisis los activos dentro de un cluster:

- Idealmente tengan un perfil de retorno similar.
- Tengan un perfil de comovimientos similar.

Medidas

Desarrollamos 3 medidas de distancia entre los activos:

- Spread
- K-S
- Omega

Medidas - Spread

Se tienen las series de retornos de dos activos x,y. La medida de spread la definimos entre cada par de activos como:

$$sd(u = y - xb) \tag{4}$$

Que es la desviación estándar del spread de los retornos salvo una constante. La intuición es que salvo una constante dos activos presentan co-movimientos similares.

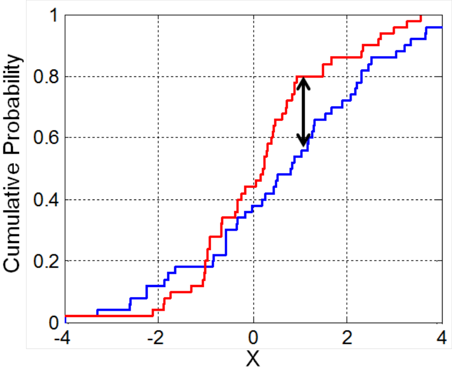
Medidas - KS

La medida K-S es el valor del estadístico de Kolmogorov-Smirnoff.

$$D_n = sup_X |F_y(X) - F_{xb}(X)| \tag{5}$$

Que es la distancia máxima entre las CDF de los retornos de los dos activos salvo una constante.

Gráficamente el estadístico K-S:



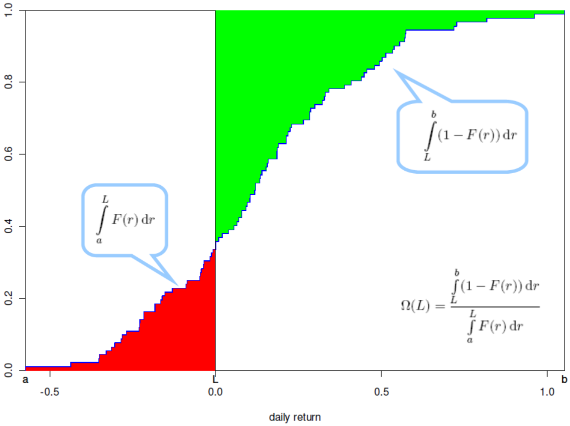
Medidas - Omega

La medida de riesgo Omega fue introducida por Keating et al. (2002) y esta definida como:

$$\Omega_r(L) = \frac{\int_L^b (1 - F(r))dr}{\int_a^L F(r)dr} \tag{6}$$

Donde $r = u = y - xb$. Que es el cociente entre una call de strike L sobre una put con el mismo strike sobre el activo u. En otras palabras el valor esperado de las ganancias por encima de un valor L sobre el valor esperado de las perdidas sobre el mismo valor.

Gráficamente la medida Omega:

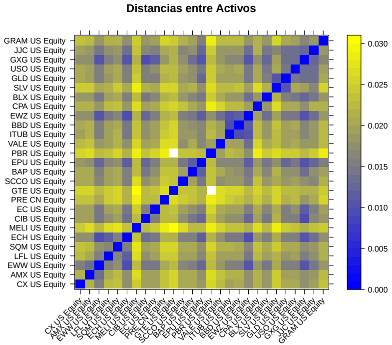


Medidas

Con las medidas anteriores es posible construir matrices de distancia entre todos los activos, donde el elemento (i, j) es justamente la distancia entre el activo i y el activo j.

Medidas

Ejemplo medida spread:



Medidas

Con estas medidas de distancias entre activos es posible aplicar una estrategia de clustering.

Contenido

k-medoides

El algoritmo de k-medoides consiste en hacer particiones de diferentes puntos en grupos minimizando la distancia entre los puntos clasificados en un mismo grupo. El algoritmo *Partitioning Around Medoids* (PAM) consiste en:

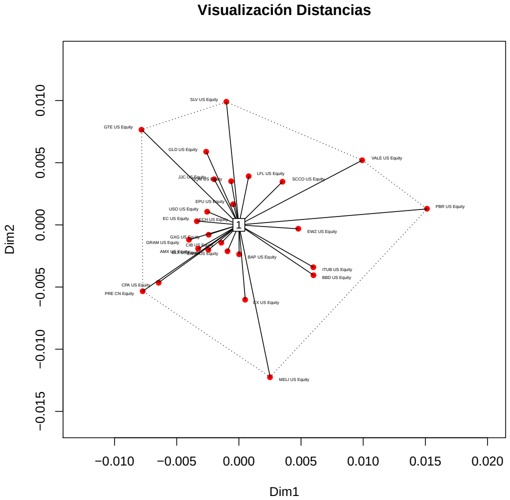
- Seleccionar aleatoriamente k puntos como medoides.
- Se asocia cada punto sobrante al medoide mas cercano.
- Para cada medoide m.
    - Para cada punto no medoide o.
        - Intercambiar m y o. Y calcular el costo total del nuevo arreglo.
- Seleccionar el arregla con el menor costo total.
- Repetir pasos 2-4 hasta tener el menor costo total.

El costo total es la suma de las distancias a su medoide correspondiente de todos los puntos.

k-medoides

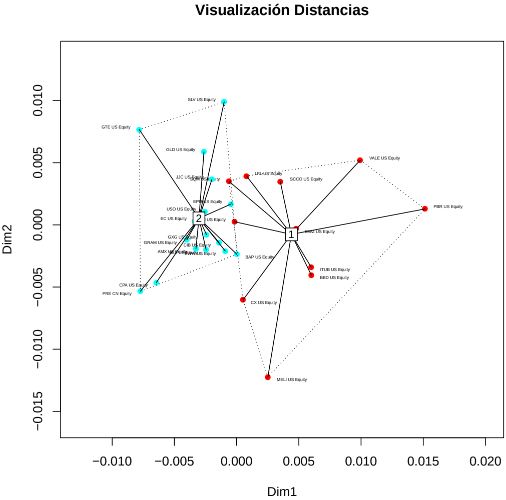
- Una de las debilidades de esta metodología es que no establece un criterio de optimalidad en el número de clusters (k).
- La sensibilidad al número de clusters (k) es evidente.
- Las siguientes representaciones gráficas de los clusters se construyen usando *Multidimensional scaling* (MDS)

k-medoides



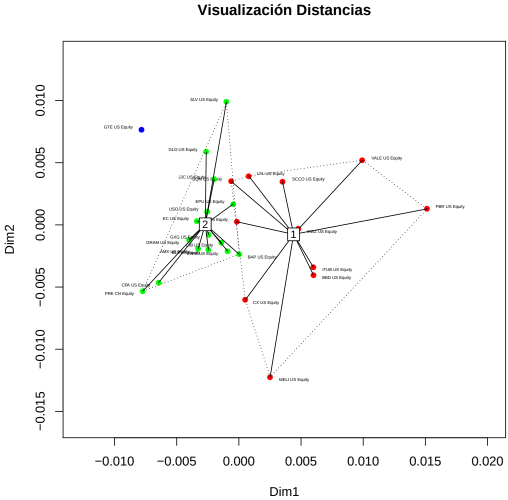
Visualización Distancias
Dim1
Dim2
−0.010
−0.005
0.000
0.005
0.010
0.015
0.020
0.010
0.005
0.000
−0.005
−0.010
−0.015
1
SLV US Equity
GTE US Equity
GLD US Equity
VALE US Equity
LFL US Equity
SCCO US Equity
JJC US Equity
SQM US Equity
PBR US Equity
EPU US Equity
USO US Equity
EC US Equity
ECH US Equity
EWZ US Equity
GXG US Equity
GRAM US Equity
CIB US Equity
AMX US Equity
BAP US Equity
ITUB US Equity
BBD US Equity
CPA US Equity
PRE CN Equity
CX US Equity
MELI US Equity

k-medoides

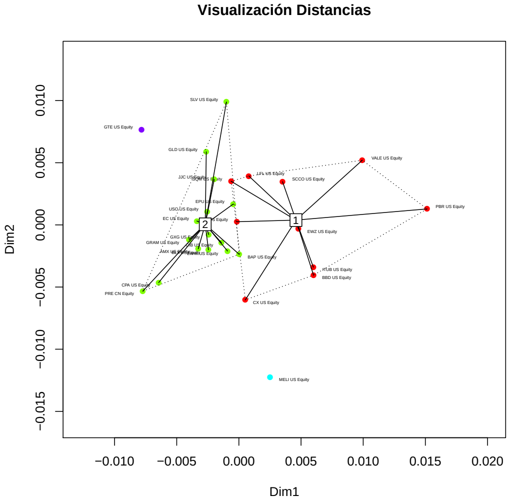


Visualización Distancias
SLV US Equity
GTE US Equity
GLD US Equity
VALE US Equity
LFL US Equity
SCCO US Equity
PBR US Equity
EPU US Equity
USO US Equity
EC US Equity
2
1
EWZ US Equity
GXG US Equity
GRAM US Equity
CIB US Equity
BAP US Equity
ITUB US Equity
BBD US Equity
CPA US Equity
PRE CN Equity
CX US Equity
MELI US Equity
Dim1
Dim2
−0.010
−0.005
0.000
0.005
0.010
0.015
0.020
−0.015

k-medoides

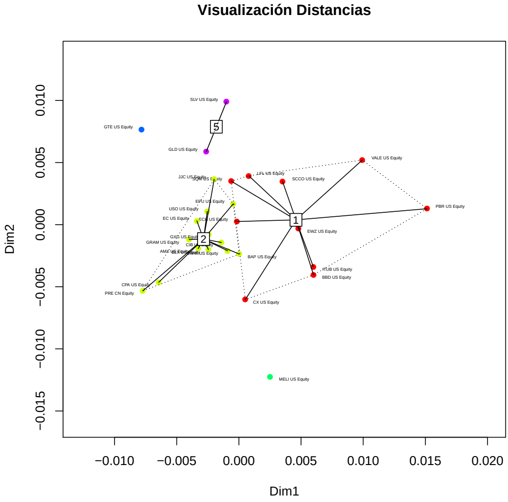

Visualización Distancias
SLV US Equity
GTE US Equity
GLD US Equity
VALE US Equity
LFL US Equity
SCCO US Equity
JJC US Equity
SQM US Equity
PBR US Equity
EPU US Equity
USO US Equity
EC US Equity
EWZ US Equity
GXG US Equity
GRAM US Equity
AMX US Equity
EWW US Equity
BAP US Equity
ITUB US Equity
BBD US Equity
CPA US Equity
PRE CN Equity
CX US Equity
MELI US Equity
1
2
0.010
0.005
0.000
−0.005
−0.010
−0.015
Dim2
−0.010
−0.005
0.000
0.005
0.010
0.015
0.020
Dim1

k-medoides

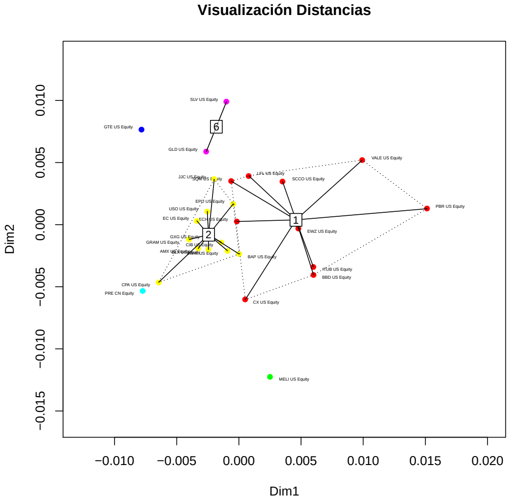


Visualización Distancias
SLV US Equity
GTE US Equity
GLD US Equity
JJC US Equity
LFL US Equity
SCCO US Equity
VALE US Equity
PBR US Equity
EPU US Equity
USO US Equity
EC US Equity
GXG US Equity
GRAM US Equity
CIB US Equity
AMX US Equity
BAP US Equity
EWZ US Equity
ITUB US Equity
BBD US Equity
CPA US Equity
PRE CN Equity
CX US Equity
MELI US Equity
1
2
Dim2
Dim1
0.010
0.005
0.000
−0.005
−0.010
−0.015
−0.010
−0.005
0.000
0.005
0.010
0.015
0.020

k-medoides

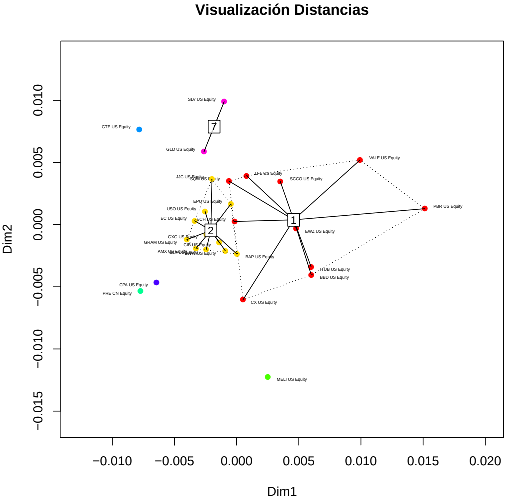

Visualización Distancias
SLV US Equity
GTE US Equity
GLD US Equity
5
VALE US Equity
LFL US Equity
SCCO US Equity
JJC US Equity
PBR US Equity
EPU US Equity
USO US Equity
EC US Equity
1
EWZ US Equity
2
GRAM US Equity
BAP US Equity
ITUB US Equity
BBD US Equity
CPA US Equity
PRE CN Equity
CX US Equity
MELI US Equity
Dim2
0.010
0.005
0.000
−0.005
−0.010
−0.015
−0.010
−0.005
0.000
0.005
0.010
0.015
0.020
Dim1

k-medoides



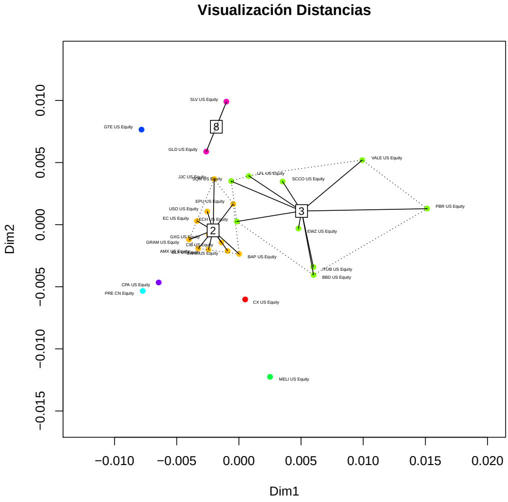
Visualización Distancias
SLV US Equity
GTE US Equity
GLD US Equity
6
VALE US Equity
LFL US Equity
SCCO US Equity
JJC US Equity
SQM US Equity
PBR US Equity
EPU US Equity
USO US Equity
EC US Equity
ECH US Equity
1
2
EWZ US Equity
GXG US Equity
GRAM US Equity
CIB US Equity
AMX US Equity
BAP US Equity
ITUB US Equity
BBD US Equity
CPA US Equity
PRE CN Equity
CX US Equity
MELI US Equity
Dim2
Dim1
0.010
0.005
0.000
−0.005
−0.010
−0.015
0.015
0.020

k-medoides



Visualización Distancias
SLV US Equity
GTE US Equity
7
GLD US Equity
VALE US Equity
LFL US Equity
SCCO US Equity
JJC US Equity
EPU US Equity
USO US Equity
EC US Equity
ECH US Equity
PBR US Equity
1
2
EWZ US Equity
GXG US Equity
GRAM US Equity
CIB US Equity
AMX US Equity
BAP US Equity
ITUB US Equity
BBD US Equity
CPA US Equity
PRE CN Equity
CX US Equity
MELI US Equity
0.010
0.005
0.000
−0.005
−0.010
−0.015
Dim2
−0.010
−0.005
0.000
0.005
0.010
0.015
0.020
Dim1

k-medoides



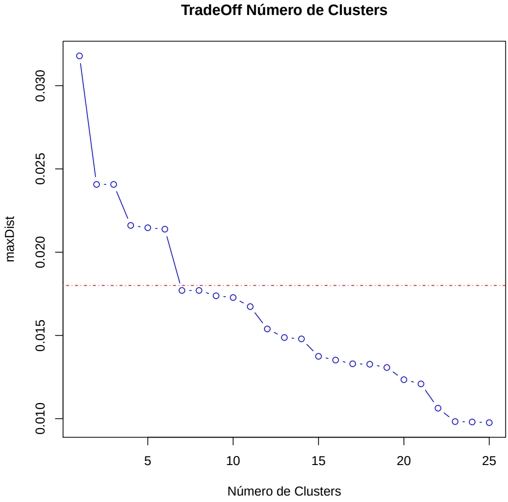
Visualización Distancias
SLV US Equity
GTE US Equity
8
GLD US Equity
VALE US Equity
LFL US Equity
JJC US Equity
SQM US Equity
SCCO US Equity
PBR US Equity
EPU US Equity
USO US Equity
3
EC US Equity
ECH US Equity
2
EWZ US Equity
GXG US Equity
GRAM US Equity
CIB US Equity
AMX US Equity
BAP US Equity
ITUB US Equity
BBD US Equity
CPA US Equity
PRE CN Equity
CX US Equity
MELI US Equity
Dim1
Dim2
−0.010
−0.005
0.000
0.005
0.010
0.015
0.020
−0.015

Número de Clusters

En un principio establecimos el siguiente criterio para determinar el número de clusters:

- Se establece como costo la máxima distancia a su medoide respectivo de todos los puntos.
- El punto donde el cambio marginal del costo sea máximo sera correspondiente al número de clusters.

Número de Clusters

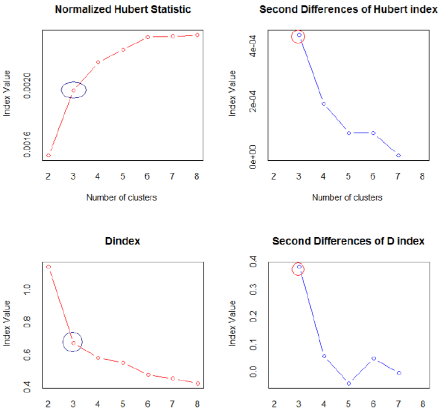


Número de Clusters

Sin embargo, Milligan y Cooper (1985) examinan mas de 40 indices para determinar el número óptimo de clusters. Por ejemplo, el indice de Dunn(1974) consiste en el cociente entre la minima distancia interclusters sobre la maxima distancia interclusters.

- Actualmente se están examinando cuales indices tienen mejor impacto en el backtest de la estrategia de selección de portafolio.

Los resultados de los diferentes indices son variados para los mismos datos.



Clustering Jerárquico

También dentro del análisis se esta usando clustering jerarquico conglomerativo, que consiste en:

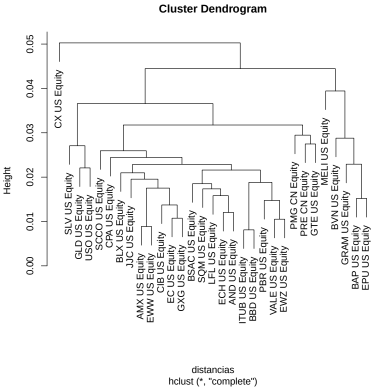
- Todos los activos inicialmente pertenecen a su propio cluster.
- Se agrupan dos clusters proximos en un mismo cluster de acuerdo a la siguiente formula para clusters X, Y:

$$D(X, Y) = \max_{x \in X, y \in Y} d(x, y) \tag{7}$$

- Continua el proceso iterativo hasta que todos los elementos estén en el mismo cluster.

Clustering Jerárquico

El dendograma del clustering jerarquico se ve así para la métrica de spread.



Contenido

Meta-heurístico de Optimización

- En la sección anterior se obtiene una partición del universo de activos en *clusters* o grupos de activos.
- La siguiente pregunta es entonces como obtener una asignación optima de activos basado en los clusters.

Meta-heurístico de Optimización

Para el problema de inversión desarrollamos el siguiente meta-heurístico:

- Se calculan los clusters dentro de una medida escogida.
- Se realiza una optimización intra-clusters.
- A partir de la optimización anterior se obtienen 2 activos representativos del cluster (no necesariamente activos individuales, sino una combinación lineal de los activos del cluster). Uno largo y uno corto.
- Con estos activos representativos se construye una frontera eficiente usando unicamente los activos representativos del cluster.

Meta-heurístico de Optimización

Optimización intra-cluster:

- Se optimiza una frontera eficiente dentro del cluster (es decir, se construye una frontera eficiente usando solamente los activos del cluster).
- Se eligen dos activos representativos (portafolios) de la siguiente manera:
  - max E[u(r)]
  - max E[u(-r)]
- La intuición es obtener dos activos por cluster. Uno con perfil *Largo* y otro con perfil *Corto*.

Optimización inter-cluster:

- Con estos activos $(2n)$, donde n es la cantidad de clusters, se construye una frontera eficiente derivada del siguiente problema:
  - $$\min w^T \Omega w \tag{8}$$

    Sujeto a:

    $$\begin{aligned} \sum_i \mu_i w_i &= \bar{r} \\ \sum_i w_i &= 1 \\ w_i^{min} \leq w_i &\leq w_i^{max} \end{aligned} \tag{9}$$
  - No se puede estar invertido simultáneamente en los dos activos de un mismo cluster.
  - Los activos largos de los clusters tienen limites mayores o iguales a 0: $0\,\% \leq w_i^L$
  - Los activos cortos de los clusters tienen limites menores o iguales a 0: $0\,\% \geq w_i^C$

Meta-heurístico de Optimización

- Finalmente se elige un portafolio que maximice la función de utilidad deseada.
- Las ventajas de la estrategia de optimización son:
  - Se hace un *hedging* más sensible a cambios no capturados por las covarianzas (dado que se agrupan los activos en clusters y no se permiten largos y cortos entre activos similares).
  - La volatilidad del portafolio debería disminuir. *Hipótesis por probar*
- Las desventajas de la estrategia de optimización son:
  - Puede ser difícil operativamente estar invertido en todos los activos simultáneamente.
  - Es necesario determinar una estrategia de tiempo de rebalanceo de los clusters.

Meta-heurístico con Cardinalidad

- Dentro del mismo problema general de optimización de portafolio se puede generalizar el meta-heurístico usando restricciones de cardinalidad.
- El problema de selección de portafolio con restricciones de cardinalidad es:

$$\min w^T \Omega w \tag{10}$$

Sujeto a:

$$\begin{aligned} \sum_i \mu_i w_i &= \bar{r} \\ \sum_i w_i &= 1 \\ \delta_i w_i^{min} \leq w_i &\leq \delta_i w_i^{max} \\ \sum_i \delta_i &= K \\ \delta_i &\in 0, 1 \end{aligned} \tag{11}$$

Meta-heurístico con Cardinalidad

- Para solucionar este problema se usa un algoritmo llamado *Branch and Bound*, propuesto por Ailsa Land y Alison Doig en 1960: 'An automatic method of solving discrete programming problems.'
- El algoritmo busca dentro de todo el universo de posibles candidatos, usando una estrategia iterativa que le permite ir descartando en masa candidatos que no son parte de la solución posible del problema.
- `http://rjlipton.wordpress.com/2012/12/19/branch-and-bound-why-does-it-work/`

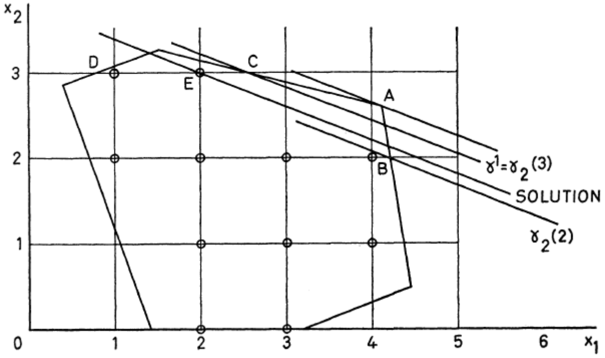
La idea detrás del paper original es la siguiente:

- Se tiene un problema de programación lineal de la siguiente forma:

$$
\begin{aligned}
c'x + \bar{c}'y &= \gamma \\
\textbf{st} \\
Ax + \bar{A}y &\leq b \\
\textbf{x es un entero no negativo} \\
y &\geq 0
\end{aligned}
\tag{12}
$$

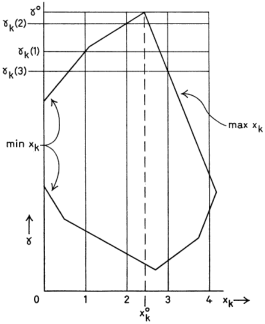
- Donde γ, el máximo, es un escalar. b es un vector columna de m filas. c y x son vectores columna de n_1 filas. $\bar{c}$ y y son vectores columna de $(n - n_1)$ filas. A es una matriz de dimensiones mxn_1 y $\bar{A}$ es una matriz de dimensiones $mx(n - n_1)$

La solución e idea detrás del paper es mover el funcional de la solución γ hasta un valor γ_0 en el cual se cumplan las todas restricciones enteras.



Meta-heurístico con Cardinalidad

El procedimiento para mover el funcional en dimensiones superiores es tomar una de las variables x y calcular el limite máximo de γ de tal forma que se cumplan las restricciones enteras, por ejemplo:



En este caso se sabe que los candidatos a la solución tienen un limite superior de $\gamma_k(2)$

Meta-heurístico con Cardinalidad

Se tiene entonces el siguiente algoritmo.

- $P(j)$: maximizar γ sujeto a las restricciones continuas y adicionalmente a j de las x variables discretas $(j = 0, 1, 2, ..., n_1)$.
- Sea S_j el conjunto de las soluciones alcanzables del tipo $P(j)$ y sea $\bar{S}$ el conjunto de las soluciones no alcanzables del tipo $P(j)$.
- El valor de S_{n1} que maximiza γ es la solución del problema.
- La solución del problema esta acotada por arriba por el valor máximo de γ en el conjunto S_{n_1-1} y así sucesivamente.
- La solución se construye haciendo un grafo de árbol donde los vértices son elementos de los conjuntos S_j.

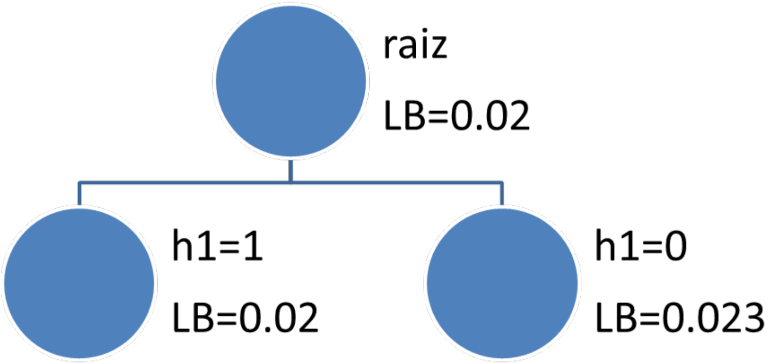
Meta-heurístico con Cardinalidad

Los pasos para construir el árbol son:

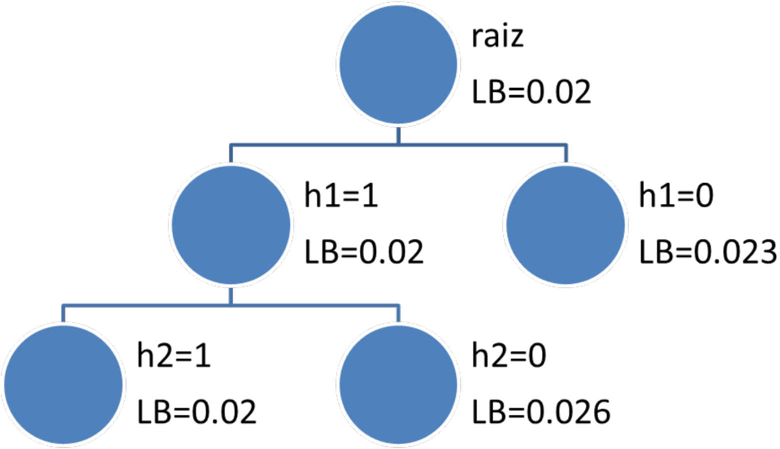
- Paso 0: El primer vértice del árbol es la solución de $P(0) = \gamma_0$, si la solución de $P(0)$ satisface $P(n_1)$ entonces se ha solucionado el problema.
- Paso 1: Si γ_0 no es la solución del problema, entonces de acuerdo a las reglas abajo, se dibujan nuevas aristas correspondientes a los puntos en S_1 y se evalúa γ en los puntos de S_1 y nunca en los de $\bar{S}$.
- Paso 2: Si los vértices $\gamma_0, ..., \gamma_k$ ya han sido rotulados con el valor γ, al vértice con el mayor γ no rotulado, se le da el rotulo γ_k.
- Paso 3: La solución final se ha alcanzado cuando por primera vez un vértice es un elemento de S_n; esto ocurre tan pronto todas las variables x son enteros no negativos. Si γ_k no es una solución de este tipo, suponga que es un elemento de S_j. Una nueva arista se origina en el vértice inmediatamente arriba de γ_k y terminando en el mismo subconjunto de soluciones (S_j) que γ_k o en $\bar{S}$. Si este nuevo vertice esta en S_j, su valor γ es necesariamente menor o igual a γ_k.
- Paso 4: Se dibujan dos vértices desde γ_k a los puntos en S_{j+1} o $\bar{S}$. De

Algoritmo de optimización B&B

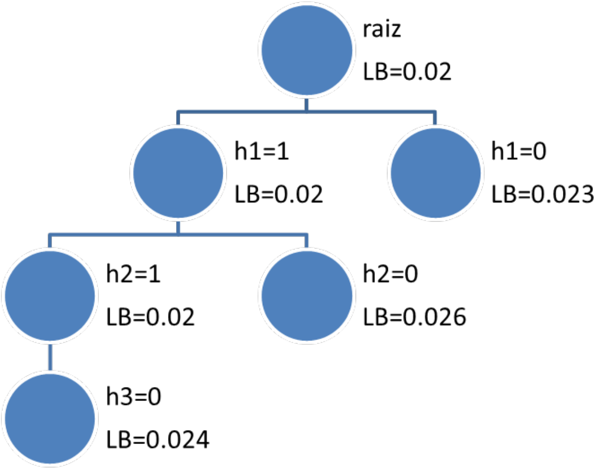
Ejemplo en el problema de Cardinalidad, se tienen 4 activos y se quiere elegir un portafolio de 2.



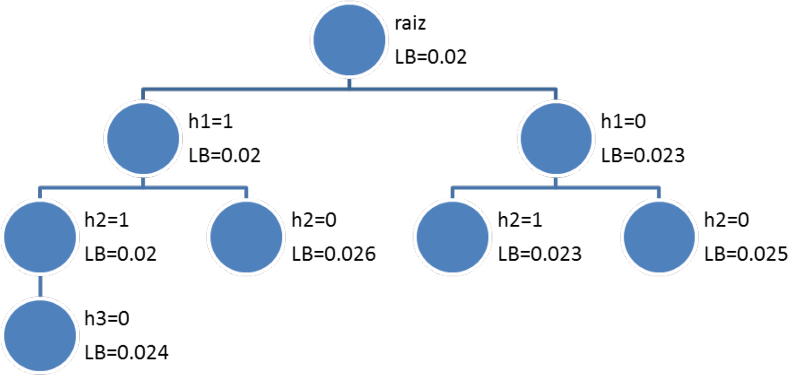
Algoritmo de optimización B&B



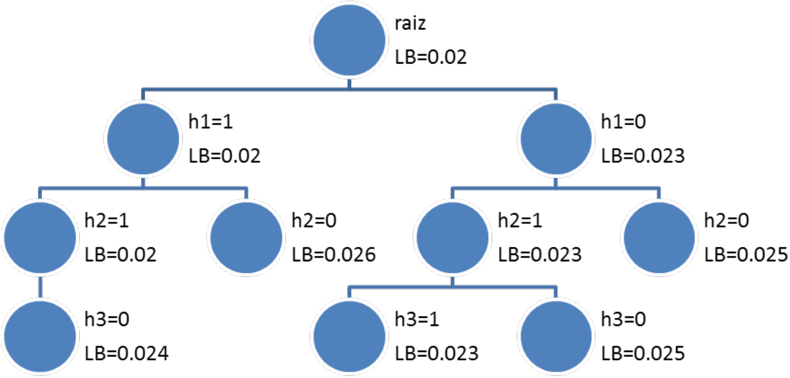
Algoritmo de optimización B&B

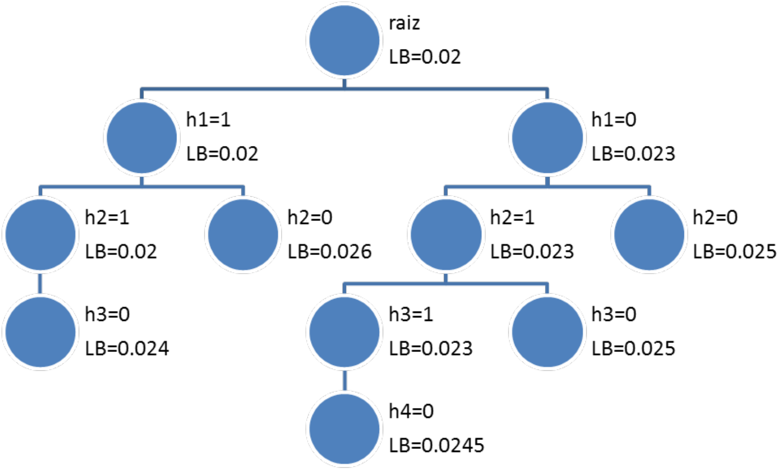


Algoritmo de optimización B&B

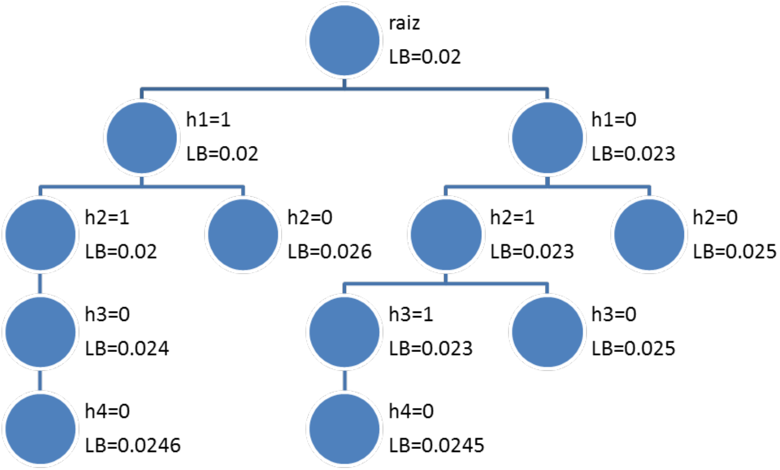


Algoritmo de optimización B&B




raiz
LB=0.02
h1=1
LB=0.02
h1=0
LB=0.023
h2=1
LB=0.02
h2=0
LB=0.026
h2=1
LB=0.023
h2=0
LB=0.025
h3=0
LB=0.024
h3=1
LB=0.023
h3=0
LB=0.025
h4=0
LB=0.0245

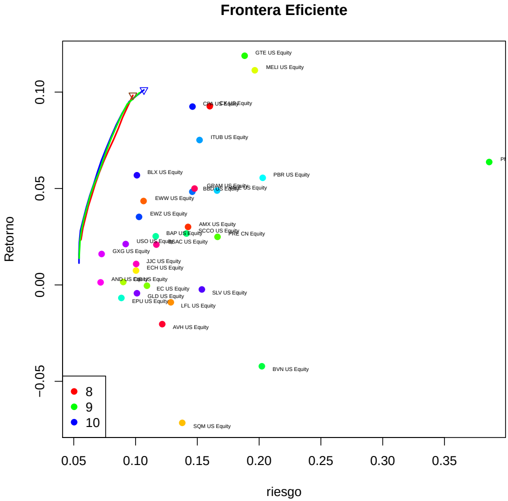
Algoritmo de optimización B&B



Algoritmo de optimización B&B

Con esto el portafolio que tiene los activos 2 y 3 es el que soluciona el problema de cardinalidad.

Una frontera eficiente con este número de activos se ve así:



Algoritmo de optimización B&B

La forma en como se conecta el problema de clusters con el de cardinalidad es usar una restricción de cardinalidad en el problema intra-clusters cuando se quiere resolver el problema usando la metodología de clusters para acotar el número total de activos que se van a tener en el portafolio.