

OPTIMIZACIÓN DE CARTERA DE ACCIONES CON PYTHON

Alejandro Bartolomeo

Profesor Titular
Matemática financiera
Inversiones Financieras

Gustavo Machín Urbay

Jefe de Trabajos Prácticos
Matemática financiera
Auxiliar de Primera
Inversiones Financieras

1. Introducción

Un **modelo de optimización** intenta determinar la mejor combinación de variables de decisión con el fin de obtener los resultados deseados.

La optimización tiene como objeto encontrar para una o más variables de decisión, cuál (es) valor (es) maximiza o minimiza una variable de desempeño. En términos matemáticos, encontrar x que optimiza una $f(x)$.

- x : variable de decisión.
- $f(x)$: variable de desempeño o variable objetivo.

El problema de la optimización de carteras de acciones puede enfocarse de dos formas. Se busca **maximizar el rendimiento, dado niveles variables de riesgo** o bien, **minimizar el riesgo dado niveles variables de rendimiento**. Este concepto, conocido como **Teoría del Conjunto Eficiente**, será desarrollado más adelante en este trabajo. Cabe aclarar que el enfoque de las inversiones en el presente trabajo de investigación es de inversiones de largo plazo (*buy and hold*).

En la actualidad las tecnologías de información desempeñan un papel muy importante en el área financiera. Uno de ellos es el uso de lenguajes de programación para la formación de una cartera de acciones optimizada. El lenguaje más versátil para lograr ese cometido es Python.

2. El Problema de la Selección de Cartera

Todas las acciones en que se puede invertir tienen resultados inciertos y por lo tanto son riesgosas. El problema básico que enfrenta cada inversionista en los mercados de capitales es determinar **qué títulos riesgosos poseer y en qué proporciones**. A esta situación se la conoce como el **problema de selección de cartera**.

Harry Markowitz presentó una solución al problema de selección de cartera en un documento (Markowitz, 1952) que se considera el origen de la **teoría moderna de la cartera**. Según esta teoría, la selección de los títulos y sus proporciones se basa en **dos supuestos** fundamentales: ante niveles equivalentes de riesgo siempre se preferirán los rendimientos más altos a los más bajos (**Insaciabilidad**) y ante el mismo rendimiento se escogerá la cartera con nivel de riesgo menor (**Aversión al riesgo**).

3. Modelos de optimización

Un **modelo de optimización** intenta determinar la mejor combinación de variables de decisión con el fin de obtener los resultados deseados.

En las inversiones bursátiles existen “riesgos” que pueden producir cuantiosas pérdidas debido fundamentalmente, a variaciones (volatilidad) del mercado. En un intento de dar respuesta al problema planteado; se propone la creación una **cartera optimizada**; es decir, la mejor combinación de acciones (variables de decisión) que entregue el menor riesgo para un máximo rendimiento de la inversión.

4. Variables o componentes del modelo de optimización

Los modelos de optimización presentan al menos los siguientes componentes:

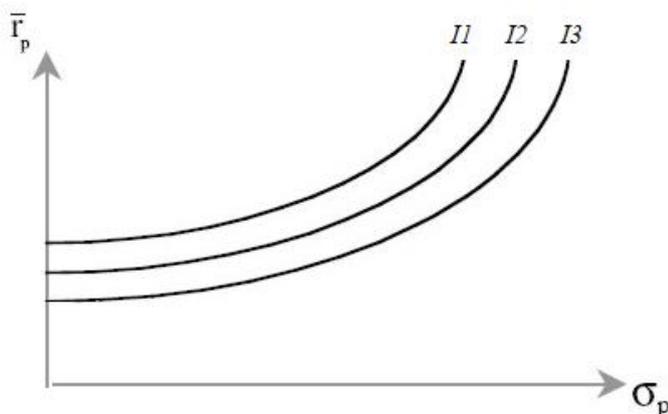
- El **objetivo** es aquello que se pretende optimizar, es decir, alcanzar el mejor valor posible. Algunos objetivos para el problema de optimización de carteras podrían ser:
 - **Maximizar el rendimiento:** se busca la combinación de acciones, que, para niveles variables de riesgo, ofrezcan el máximo rendimiento esperado.
 - **Minimizar el riesgo:** se busca la combinación de acciones, que, para niveles variables de rendimiento, ofrezcan el mínimo riesgo.
- Las **variables de decisión** son los valores controlables. Las variables de decisión del modelo serán la participación (peso relativo) de cada una de las acciones (k_1, k_2, \dots, k_n), en la cartera del inversor.
- Las **restricciones** son limitaciones en los valores que pueden tomar las variables de decisión. En un problema de optimización de cartera de acciones, las restricciones son principalmente: las participaciones de las acciones en la cartera deben ser mayor o igual a cero y a su vez deben ser complementarias, es decir, deben sumar uno.

5. ¿Qué optimizamos, la dispersión o el rendimiento?

El problema de la **optimización de carteras de acciones** puede enfocarse de dos formas. Se busca **maximizar el rendimiento**, dado un nivel de riesgo o bien **minimizar el riesgo** dado un nivel de rendimiento. Si se da el nivel de riesgo en el primer caso o el de rendimiento en el segundo, lo que se está haciendo, de alguna forma, es fijar cuál es la **curva de indiferencia del inversor**.

Las **curvas de indiferencia** representan todas las combinaciones posibles de riesgo y rendimiento que proporcionan la misma utilidad a un inversor. Por lo tanto, este inversor es indiferente a las combinaciones rendimiento-riesgo de la misma curva de indiferencia. Las curvas de indiferencia se pueden dibujar en una figura bidimensional donde el eje horizontal indica el riesgo medido por la desviación estándar y el eje vertical indica el rendimiento esperado. (Alexander et al, 2003).

Gráfico N° 1
Curvas de indiferencia



5.1. Rendimiento

Antes de hablar del rendimiento de las acciones, será necesario realizar algunas consideraciones previas sobre su cálculo, ya sea diario, mensual o en base a otro período de tiempo.

El rendimiento periódico (de un día, de un mes, de un año o en base a otro período de tiempo) de una acción puede calcularse como:

$$R_{(n-j;n)} = \frac{P_n}{P_{n-j}} - 1 \quad (1)$$

Donde:

- $R_{(n-j;n)}$: rendimiento del título en el período n (entre los momentos $n-j$ y n).
- P_n : precio del título en el período n .
- P_{n-j} : precio del título en el período $n-j$ (j períodos antes).

En general podría ocurrir, que los precios sobre los que se calculan los rendimientos, tengan cierta variabilidad. Ésta tiene consecuencias, dependiendo del modo en que calculemos el rendimiento con respecto al período deseado. Para comenzar, podríamos decir que los rendimientos podrían calcularse como una media aritmética, para la unidad de tiempo j (día, mes, año).

No obstante, este último planteo no siempre arroja resultados razonablemente representativos. Veamos un ejemplo:

| Momento | Cotización | Rendimiento [0;1] y [0;2] | Rendimiento [0;2] |
|---------|------------|--------------------------------|-------------------|
| 0 | 100 | | |
| 1 | 50 | $\frac{50}{100}-1=-0,50=-50\%$ | 100/100-1=0% |
| 2 | 100 | $\frac{100}{50}-1=1=+100\%$ | |

La **media aritmética** entre los rendimientos, teniendo en cuenta los resultados de la tercer columna, es: $(-50\%+100\%)/2 = 25\%$. Sin embargo, considerando la cuarta columna, elrendimiento del período entre 0 y 2 es 0%. Esta última expresión tiene más que ver con la realidad de los precios presentados en el ejemplo, ya que los precios no sufrieron cambios desde 0 a 2. Sin embargo, el primer procedimiento muestra un rendimiento promedio del 25%.

Para solucionar estas dificultades se pueden plantear dos métodos alternativos para promediar los rendimientos. Esto son la **media logarítmica** y la **media geométrica**.

La **media logarítmica** usa los logaritmos naturales para el cálculo de los rendimientos y luego la media aritmética. Para el ejemplo sería: $[\ln(\frac{50}{100}) + \ln(\frac{100}{50})] / 2 = 0$

La **media geométrica** se define como la raíz n del producto de n rendimientos menos uno. Para el ejemplo anterior sería: $\sqrt[2]{2 \times 0,5} - 1 = 0$

Una vez hechas estas aclaraciones, es necesario definir qué se entiende por **rendimiento esperado (R)**.¹ Existen diversos métodos para calcular los rendimientos esperados de activos financieros (los que plantean el descuento de dividendos, los que se basan en relaciones de precio y ganancia, y los modelos de CAPM). (Elbaum, 2004).

Los métodos que se basan en cuál será el dividendo o ingreso esperado de la acción descuentan esos flujos a una tasa que refleje no solo el valor del dinero en el tiempo (o recompensa por espera) sino también el riesgo asumido. Esta expectativa, de alguna manera se traducirá, más tarde o más temprano en el precio del activo hoy. Por lo tanto, bajo la hipótesis de mercados eficientes, en donde la información fluye, llega a todos los inversores y se refleja inmediatamente en los precios de las acciones, el rendimiento esperado en el futuro es el que se deriva de los precios de los activos hoy. Bajo estas condiciones, no sería disparatado suponer que los precios del futuro podrían comportarse como lo han hecho en el pasado, ya que han reflejado toda la información existente respecto del activo en el precio en ese momento. Los cambios que se han producido en los precios se han dado **aleatoriamente**. Ya sean estos buenos o malos, se han producido sorpresivamente y se han reflejado en los precios tan rápido como sea posible.

¹ Se puede definir el **rendimiento esperado**, como el rendimiento de una acción que el inversionista **espera** recibir durante el período de tenencia, lapso de tiempo durante el cual invertirá su dinero.

Más allá del método elegido, cabe preguntarse qué escenarios aleatorios se esperan para el futuro. El rendimiento esperado en el caso más sencillo en que todos los probables rendimientos de cierta inversión financiera tengan la misma probabilidad de ocurrir, es el promedio aritmético de los rendimientos individuales de cada uno de los títulos valores que intervienen en dicha inversión.

$$\bar{R}_i = \frac{1}{n}(R_1 + R_2 + \dots + R_n) \quad (2)$$

Por otro lado, y en el caso más general, es posible conocer el rendimiento esperado de una inversión i cuando los n valores de la variable no son igualmente probables. Se consideralos rendimientos como una variable aleatoria con cierta distribución y lo que es necesario definir es un parámetro que recibe el nombre de **esperanza matemática, valor esperado o media de la variable aleatoria**. Entonces:

$$E(R_i) = \bar{R}_i = \sum_{t=1}^n R_{it} p_{it} \quad (3)$$

Donde:

- R_{it} : rendimiento del activo i en el evento t .
- p_{it} : probabilidad de que se produzca el evento t para el activo i .

Con esta fórmula se cubre desde el caso más sencillo (eventos equiprobables) hasta el más complejo, definiendo a la esperanza matemática de la variable aleatoria, como el momento de orden uno.

Si se incluyen las probabilidades de que se produzcan ciertos eventos, o dicho de otra forma, teniendo en cuenta la función de distribución de los rendimientos que corresponda, “... se consigue determinar un parámetro que permite establecer el más probable de los rendimientos de una inversión en condiciones de riesgo, pues para su cálculo se tiene en cuenta no solo los distintos rendimientos aleatorios posibles, sino también el grado de probabilidad con que pueden presentarse las distintas circunstancias que originarán esos rendimientos”. (Messuti, 1994, p. 179).

Por último, dada una constante k , la esperanza matemática de la constante por la variable aleatoria es igual a la constante por la esperanza de la variable. Por otro lado, la esperanza matemática de una suma de variables aleatorias es igual a la suma de las esperanzas de cada una de esas variables. Estas dos propiedades pueden resumirse al mismo tiempo, en la siguiente expresión:

$$E(k_1 R_1 + k_2 R_2 + \dots + k_n R_n) = k_1 E(R_1) + k_2 E(R_2) + \dots + k_n E(R_n) \quad (4)$$

Un caso particular para la aplicación de esta ecuación es la determinación del **rendimiento de una cartera de acciones**, que tiene un diferente grado de participación. Supongamos una cartera de tres títulos (1, 2 y 3) con una participación de $k_1=50\%$, $k_2=20\%$ y $k_3=30\%$. Si los rendimientos de los activos son: $R_1=13\%$; $R_2=20\%$ y $R_3=17\%$; el rendimiento promediado de esta cartera aplicando la ecuación (4)², será:

$$R_p = E(0,50 \times 13\% + 0,20 \times 20\% + 0,30 \times 17\%) = 0,50 \times 13\% + 0,20 \times 20\% + 0,30 \times 17\% = 15,6\%$$

Recordemos que la $E(R_1)$ es 13%, $E(R_2)$ es 20% y $E(R_3)$ es 17%; ya que se ha definido $p_{it}=1$. Si existiera una variable aleatoria definida, deberíamos haber tomado su esperanza matemática, tal cual se define en la ecuación (3).

5.2. Riesgo

Se suele definir al **riesgo** de una inversión como la probabilidad de que esa inversión difiera del **rendimiento**

² Se considera $p_{it}=1$; es decir que existe un solo estado posible del Mercado, en donde los rendimientos son lo señalados. En la simulación estocástica, esto variará según la distribución de probabilidades de cada acción.

esperado. Si la probabilidad de que una inversión difiera de su rendimiento esperado es alta, decimos que estamos ante una inversión relativamente riesgosa. Existen en la disciplina estadística, herramientas que permiten medir tal variabilidad o volatilidad. Una de estas herramientas es el parámetro llamado **varianza**. De un modo general, podemos definir a la varianza (σ^2) con la siguiente expresión:

$$\sigma^2(R_i) = \sigma^2 = \sum_{t=1}^n (R_{it} - \bar{R}_i)^2 p_{it} \quad (5)$$

Nótese que se consideran n estados posibles de la variable (desde $t=1$ hasta n). Para cada uno de esos estados, y para cada uno de las acciones incluidas en la cartera del inversor, también se ha definido una probabilidad p_{it} .

Simplificando la expresión anterior y suponiendo la formación de una cartera de dos acciones con las siguientes participaciones: el título 1 igual a k_1 y el título 2 igual a k_2 , siendo $k_1 + k_2 = 1$; es posible reescribirla del siguiente modo (considerando la ecuación 4):

$$\sigma^2 = \sum_{t=1}^n (R_{it} - \bar{R}_i)^2 p_{it} = \sum_{t=1}^n [(k_1 R_{1t} + k_2 R_{2t}) - (k_1 \bar{R}_{1t} + k_2 \bar{R}_{2t})]^2 p_{it} \quad (6)$$

Agrupando de forma diferente:

$$\sigma^2 = \sum_{t=1}^n [k_1 (R_{1t} - \bar{R}_{1t}) + k_2 (R_{2t} - \bar{R}_{2t})]^2 p_{it} \quad (7)$$

Desarrollando el cuadrado de un binomio y distribuyendo el factor p_{it} :

$$\sigma^2 = \sum_{t=1}^n [k_1^2 (R_{1t} - \bar{R}_{1t})^2 p_{1t} + k_2^2 (R_{2t} - \bar{R}_{2t})^2 p_{2t} + 2k_1 k_2 (R_{1t} - \bar{R}_{1t})(R_{2t} - \bar{R}_{2t}) p_{it}] \quad (8)$$

Dentro de la sumatoria se pueden apreciar tres términos. Los dos primeros sumandos equivalen al cuadrado de la participación de cada título multiplicado por la varianza del rendimiento de ese título (ya que al cuadrado de los desvíos se los multiplica por la probabilidad de ocurrencia). Respecto al tercer sumando, a las participaciones en el título 1 y 2 los multiplicamos por lo que se define estadísticamente como la **covarianza de los rendimientos** del título 1 con el 2 y por lógica combinatoria, del título 2 con el 1. Esa es la razón por la que dicho tercer sumando aparece multiplicado por 2.

La medida estadística conocida como **correlación** está estrechamente relacionada con la **covarianza**. De hecho, el coeficiente de correlación entre dos acciones es el cociente de la covarianza entre las dos acciones y el producto de sus desvíos estándar:

$$\rho_{ij} = \frac{\sigma_{ij}}{\sigma_i \sigma_j}$$

Donde:

- ρ_{ij} : coeficiente de correlación de los rendimientos entre la acción i y la acción j .
- σ_{ij} : covarianza de los rendimientos entre la acción i y la acción j .

Por lo tanto, el coeficiente de correlación entre las acciones es determinante del riesgo de una cartera.

5.3. Qué pasa para n acciones con el riesgo

Por último, para k_1, k_2, \dots, k_n participaciones, siendo σ_{ij} la covarianza entre la acción i y la acción j , la última expresión de la varianza de una cartera que consta de n acciones, podría calcularse como una doble sumatoria, de la siguiente forma:

$$\sigma^2 = \sum_{i=1}^n \sum_{j=1}^n k_i k_j \sigma_{ij} \quad (9)$$

Donde:

- σ_{ij} : covarianza de los rendimientos entre la acción i y la acción j .
- k_i : participación porcentual de la acción i dentro de la cartera.
- k_j : participación porcentual de la acción j dentro de la cartera.
- n : número de acciones que componen la cartera.

En definitiva, esta es la función a optimizar para un escenario donde se considera un solo estado de la naturaleza posible. Dicho de otro modo, es la función a minimizar (riesgo mínimo) para una optimización determinista, dado un cierto nivel deseado (y posible) de rendimiento de la cartera.

Si se generalizan las expresiones para n acciones, es posible escribir las expresiones de las fórmulas (8) y (9) (riesgo de la cartera) con notación matricial.

Se puede simbolizar a las participaciones en cada acción como:

$$K = \begin{bmatrix} k_1 \\ k_2 \\ \vdots \\ k_n \end{bmatrix}$$

Matriz de varianzas y covarianzas:

$$V = \begin{bmatrix} \sigma_1^2 & \sigma_{12} & \cdots & \sigma_{1n} \\ \sigma_{21} & \sigma_2^2 & \cdots & \sigma_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{n1} & \sigma_{n2} & \cdots & \sigma_n^2 \end{bmatrix}$$

Vector traspuesto de las k participaciones:

$$K^T = [k_1 \quad k_2 \quad \cdots \quad k_n]$$

En virtud de lo expresado, puede definirse el riesgo de la cartera para n participaciones, en forma abreviada, de acuerdo a la siguiente notación matricial:

$$\text{Riesgo de la cartera} = \sigma^2 = K^T V K$$

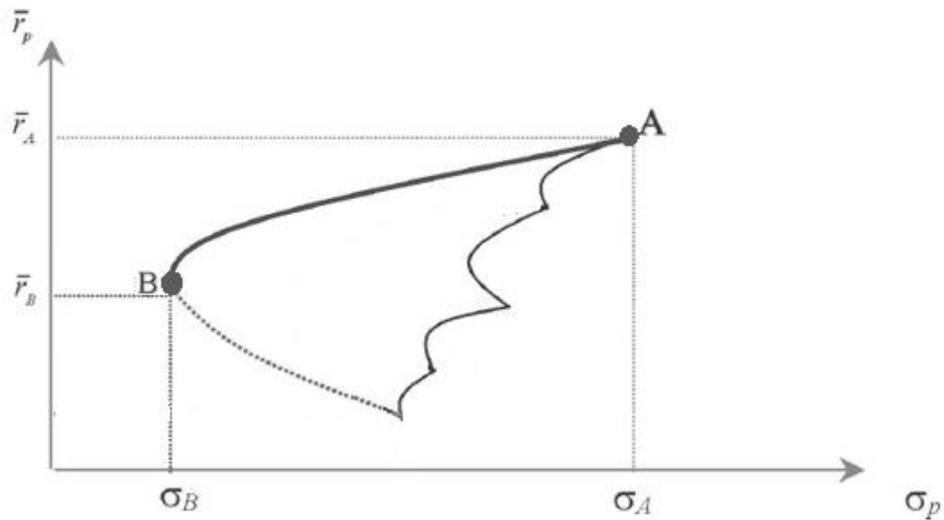
5.4. Frontera eficiente y curvas de indiferencia

Si se combinan los diferentes niveles de riesgo y rendimiento de una cartera, podrían formarse un número infinito de carteras, a partir de un conjunto de N valores. Pero el inversionista no necesita evaluar todas esas carteras. Solo buscará un subconjunto de carteras en función de lo que determina el **teorema del conjunto eficiente**. Éste establece que un inversionista elegirá su cartera óptima del conjunto de carteras que:

1. Ofrezcan el máximo rendimiento esperado para niveles variables de riesgo y,
2. ofrezcan un mínimo riesgo para niveles variables de rendimiento esperado.

El conjunto de carteras que cumplen estas 2 condiciones, se conoce como conjunto eficiente o frontera eficiente.

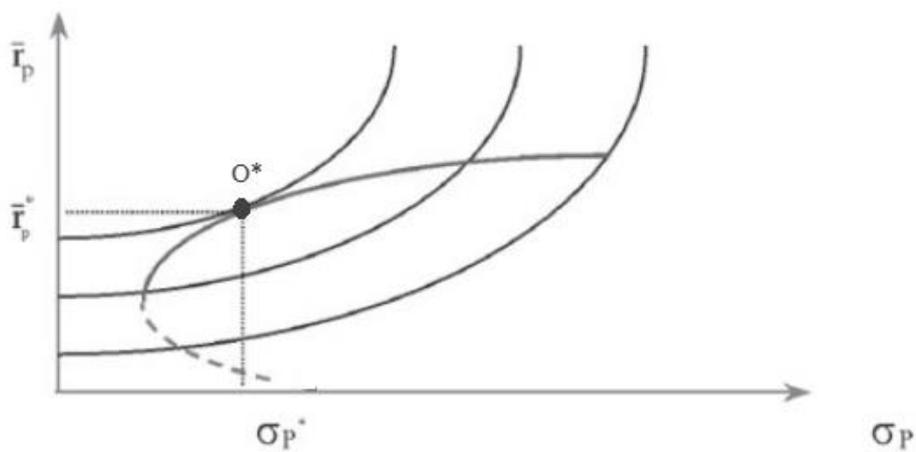
Gráfico N° 2
Conjunto eficiente



La línea marcada que se encuentra entre los dos puntos A y B (cartera de máximo rendimiento y de mínimo riesgo respectivamente) representa la frontera eficiente.

Finalmente, el inversionista debe expresar sus curvas de indiferencia de la misma forma que la frontera eficiente y luego escoger la cartera que sea tangente al conjunto eficiente.

Gráfico N° 3
Cartera óptima



5.5. Índice de Sharpe

Este índice fue desarrollado por el nobel de economía William Sharpe. Relaciona el rendimiento y riesgo de una inversión bursátil. En definitiva, representa el rendimiento por unidad de riesgo.

Su popularidad en las finanzas de mercado, no solo se debe a su utilidad sino también a su simpleza de cálculo. El índice se calcula dividiendo el rendimiento de una cartera menos la tasa libre de riesgo (por ejemplo, la de los Bonos de Tesoro Americano a 10 años) entre el riesgo de la cartera para el mismo período.

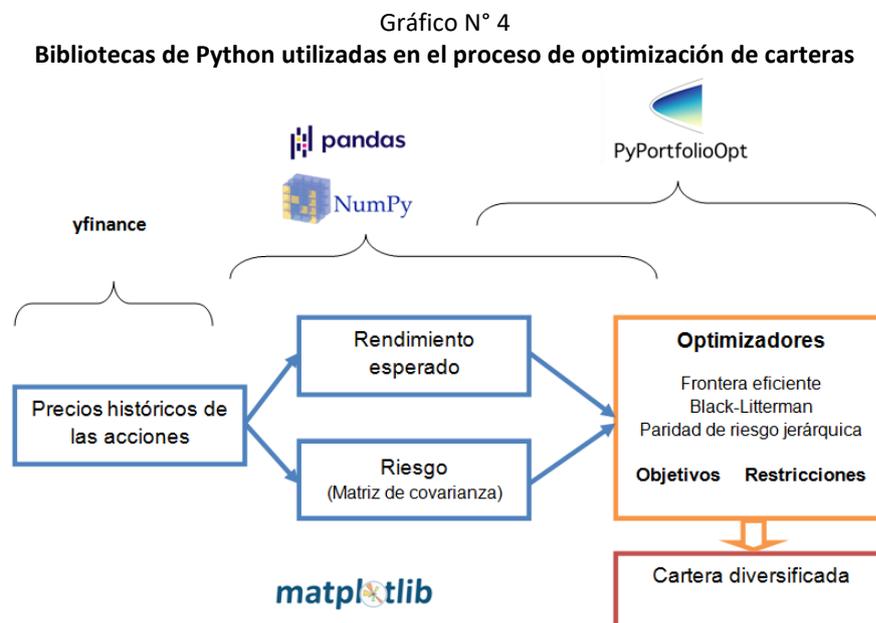
$$\text{índice de Sharpe} = \frac{R_p - r_f}{\sigma_p}$$

Donde:

- R_p : rendimiento de la cartera.
- r_f : tasa libre de riesgo.
- σ_p : riesgo de la cartera.

6. Aplicación de tecnologías a un problema de optimización de cartera

En este apartado se desarrollan paso a paso las tecnologías utilizadas para para resolver un problema de optimización de carteras, desde la obtención de los datos hasta las visualizaciones. El siguiente gráfico muestra a modo de resumen los pasos que se seguirán a continuación.



6.1. Obtención de datos financieros con *yfinance*

yfinance es una biblioteca de Python que permite obtener de manera remota los datos históricos (cotizaciones) del mercado financiero mediante la conexión a la API pública de Yahoo! Finance. Luego transforma los datos obtenidos en un *data frame* (aunque este anglicismo no tiene una traducción exacta, puede ser traducido como marco de datos). Esto lo hace de una manera muy simple, con unas poquísimas líneas de código.

yfinance no solo permite descargar los datos históricos como precios, volumen negociado o pago de dividendos de diversos activos financieros tales como acciones, bonos, criptomonedas, etc., sino que también permite descargar y visualizar todo tipo de información relevante que se encuentre disponible en Yahoo! Finance.

En este punto, solamente se utilizará la herramienta para la obtención de los precios de las acciones.

Primero se instala la librería y se importa. Posteriormente se define una lista con los símbolos (*tickers*) de los títulos que se consideran para formar la cartera. A modo de ejemplo, se tomaron las empresas tecnológicas más grandes con cotización en la bolsa de Estados Unidos.

```

pip install yfinance
import yfinance as yf
simbolos = ["AAPL", "MSFT", "TSLA", "FB", "NVDA"]
df = yf.download(simbolos, start='2021-01-01', end='2021-12-31', rounding=True)
precios = df['Adj Close']
precios

```

La función `download` se conecta a la página y descarga: el volumen negociado, los precios de apertura, cierre y cierre ajustado diarios para las acciones antes definidas y según los parámetros establecidos. Luego, se selecciona la columna correspondiente a precio de cierre ajustado (*Adjusted Close*) y se define un nuevo *data frame* (llamado *precios*).

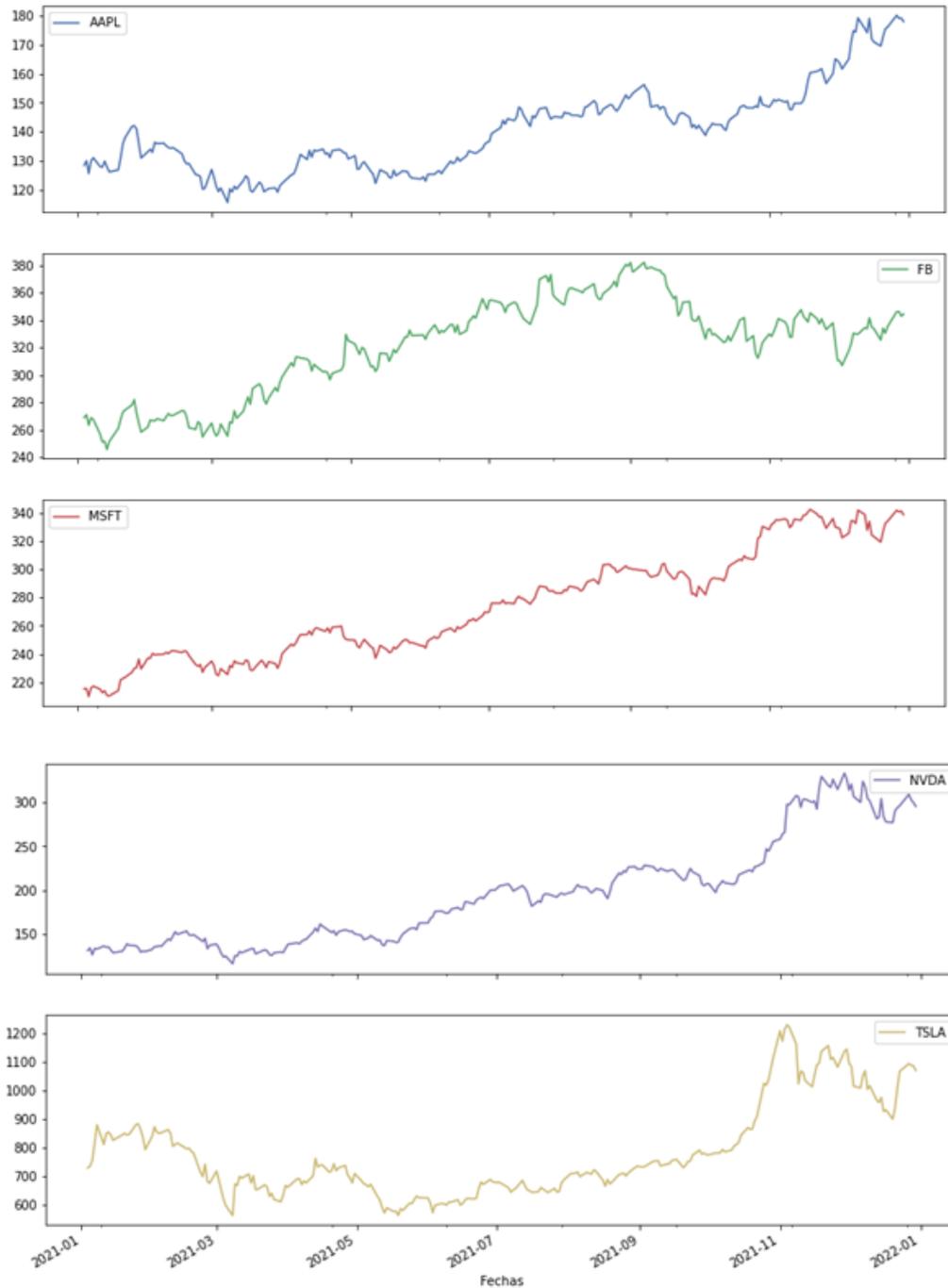
Tabla N° 1
Precio de las acciones

| | AAPL | FB | MSFT | NVDA | TSLA |
|------------|--------|--------|--------|--------|---------|
| Date | | | | | |
| 2021-01-04 | 128.45 | 268.94 | 215.43 | 131.00 | 729.77 |
| 2021-01-05 | 130.04 | 270.97 | 215.64 | 133.91 | 735.11 |
| 2021-01-06 | 125.66 | 263.31 | 210.05 | 126.01 | 755.98 |
| 2021-01-07 | 129.95 | 268.74 | 216.03 | 133.30 | 816.04 |
| 2021-01-08 | 131.07 | 267.57 | 217.34 | 132.63 | 880.02 |
| ... | ... | ... | ... | ... | ... |
| 2021-12-23 | 176.06 | 335.24 | 334.00 | 296.35 | 1067.00 |
| 2021-12-27 | 180.10 | 346.18 | 341.74 | 309.40 | 1093.94 |
| 2021-12-28 | 179.06 | 346.22 | 340.55 | 303.17 | 1088.47 |
| 2021-12-29 | 179.15 | 342.94 | 341.24 | 299.96 | 1086.19 |
| 2021-12-30 | 177.97 | 344.36 | 338.62 | 295.81 | 1070.34 |

251 rows x 5 columns

A continuación, se visualizan las cotizaciones de cada una de las acciones. Se usa la biblioteca de visualización *matplotlib*. Es posible usar otras bibliotecas tales como *plotly*, *altair* o *seaborn*. Es objetivo es realizar un análisis preliminar de cada activo.

Gráfico N° 5
Precio de las acciones



6.2. Calcular el rendimiento esperado y riesgo de cada acción

Como se explicó en la introducción teórica, los modelos de optimización requieren conocer previamente los rendimientos esperados de las acciones y la matriz de varianzas y covarianzas. En la práctica, se calculan en función de datos históricos.

La función `pct_change` de la biblioteca `pandas` calcula el cambio porcentual entre un elemento y el anterior.

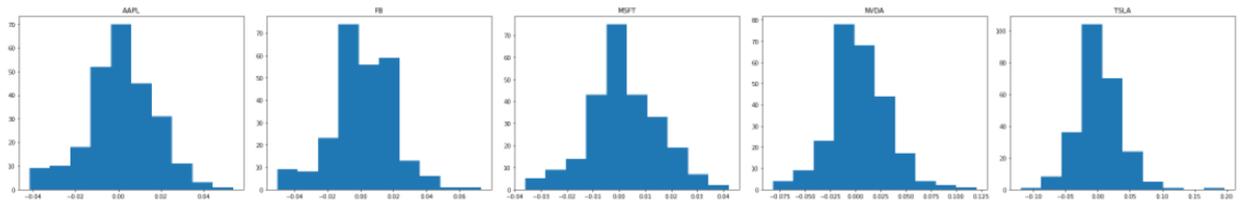
```
rendimientos = precios.pct_change()[1:]  
rendimientos
```

Tabla N° 2
Rendimiento de las acciones

| | AAPL | FB | MSFT | NVDA | TSLA |
|------------|-----------|-----------|-----------|-----------|-----------|
| Date | | | | | |
| 2021-01-05 | 0.012378 | 0.007548 | 0.000975 | 0.022214 | 0.007317 |
| 2021-01-06 | -0.033682 | -0.028269 | -0.025923 | -0.058995 | 0.028390 |
| 2021-01-07 | 0.034140 | 0.020622 | 0.028469 | 0.057853 | 0.079447 |
| 2021-01-08 | 0.008619 | -0.004354 | 0.006064 | -0.005026 | 0.078403 |
| 2021-01-11 | -0.023194 | -0.040102 | -0.009662 | 0.025937 | -0.078214 |
| ... | ... | ... | ... | ... | ... |
| 2021-12-23 | 0.003648 | 0.014495 | 0.004481 | 0.008165 | 0.057619 |
| 2021-12-27 | 0.022947 | 0.032633 | 0.023174 | 0.044036 | 0.025248 |
| 2021-12-28 | -0.005775 | 0.000116 | -0.003482 | -0.020136 | -0.005000 |
| 2021-12-29 | 0.000503 | -0.009474 | 0.002026 | -0.010588 | -0.002095 |
| 2021-12-30 | -0.006587 | 0.004141 | -0.007678 | -0.013835 | -0.014592 |

250 rows x 5 columns

Gráfico N° 6
Histograma Rendimientos



`rendimientos.describe()`

La función `describe` de Pandas genera un resumen de los estadísticos de centralización y dispersión más importantes de los datos (rendimientos) tales como media, desviación estándar y percentiles.

Tabla N° 3
Estadísticos de las acciones

| | AAPL | FB | MSFT | NVDA | TSLA |
|--------------|------------|------------|------------|------------|------------|
| count | 250.000000 | 250.000000 | 250.000000 | 250.000000 | 250.000000 |
| mean | 0.001429 | 0.001162 | 0.001897 | 0.003663 | 0.002120 |
| std | 0.015786 | 0.018606 | 0.013213 | 0.028448 | 0.034560 |
| min | -0.041674 | -0.050515 | -0.036208 | -0.082160 | -0.119903 |
| 25% | -0.007594 | -0.010438 | -0.005324 | -0.012155 | -0.015794 |
| 50% | 0.001463 | 0.000210 | 0.000966 | 0.002654 | 0.001787 |
| 75% | 0.012487 | 0.014011 | 0.010429 | 0.020280 | 0.019111 |
| max | 0.053841 | 0.072973 | 0.042117 | 0.120421 | 0.196412 |

En la tabla anterior se puede observar el rendimiento esperado (*mean*) y el riesgo (*std*) diarios de cada una de las acciones.

Si bien, es posible calcular el riesgo y rendimiento esperado de las acciones con *pandas* o *numpy*, así como armar y graficar la matriz de varianzas y covarianzas, la cual es necesaria a posteriori para calcular el riesgo de la cartera, se utilizará en este punto las funciones que provee La biblioteca *PyPortfolioOpt*.

El módulo `expected_returns` proporciona las funciones para estimar los rendimientos esperados de las acciones.

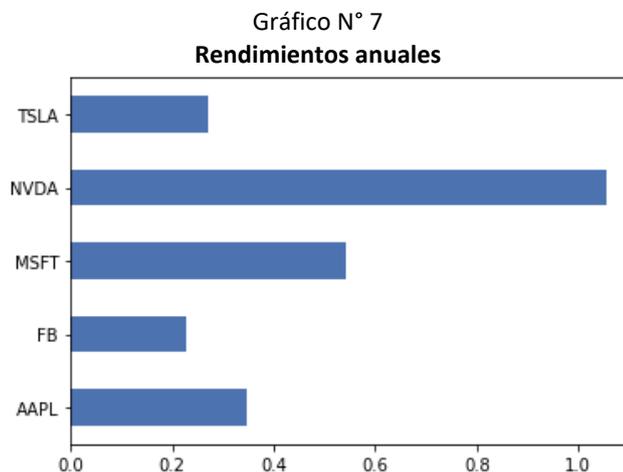
Todas las funciones internamente transforman los precios de las acciones en rendimientos porcentuales antes de calcular sus respectivas estimaciones de rendimientos esperados. Los métodos disponibles para estimar el rendimiento esperado son: rendimiento histórico medio, rendimiento histórico medio ponderado exponencialmente y CAPM. En la posterior optimización se utilizará el rendimiento esperado medio.

Por defecto, la salida de estos métodos son rendimientos esperados anuales.

```
rendEsp1 = expected_returns.mean_historical_return(precios, returns_data=False, compounding=True, frequency=252, log_returns=True)
rendEsp1

AAPL    0.346154
FB      0.228385
MSFT    0.542895
NVDA    1.055190
TSLA    0.271273
dtype: float64
```

En los parámetros de la función, se toman los precios de las acciones (no los rendimientos, tal como lo indica el parámetro `returns_data`), la frecuencia son 252 días hábiles en un año. Por último, en función de lo explicado en el punto 5.1. del presente trabajo, se utilizó la media logarítmica e interés compuesto para calcular los rendimientos anuales de las acciones.



La matriz de covarianzas, así como la matriz de correlaciones, permite observar cómo se comporta cada acción con respecto a las demás. Es posible usar la función `cov` que vienen en la biblioteca *numpy* para construir la matriz de varianzas y covarianzas.

Sin embargo, el módulo `risk_models` de la biblioteca *PyPortfolioOpt* proporciona las funciones para estimar el riesgo de las acciones. El método más usado es el que estima la matriz de varianzas y covarianzas en función de los datos históricos. Esta matriz es una representación del riesgo de las acciones individuales (diagonal) y de la covariación entre acciones.

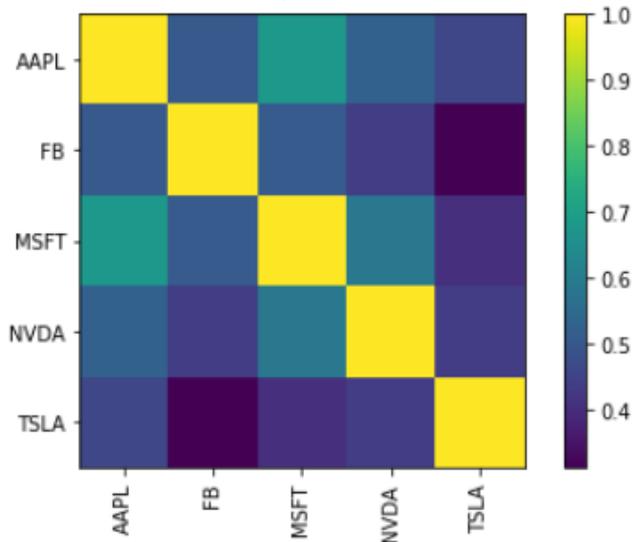
```
matrizCov1 = risk_models.sample_cov(precios)
matrizCov1
```

Tabla N° 4
Matriz de varianzas y covarianzas

| | AAPL | FB | MSFT | NVDA | TSLA |
|------|----------|----------|----------|----------|----------|
| AAPL | 0.062797 | 0.037205 | 0.035700 | 0.059513 | 0.062921 |
| FB | 0.037205 | 0.087240 | 0.031593 | 0.058230 | 0.050430 |
| MSFT | 0.035700 | 0.031593 | 0.043998 | 0.055726 | 0.046758 |
| NVDA | 0.059513 | 0.058230 | 0.055726 | 0.203936 | 0.108386 |
| TSLA | 0.062921 | 0.050430 | 0.046758 | 0.108386 | 0.300992 |

El módulo `risk_models` de *PyPortfolioOpt* permite graficar mediante un mapa de calor una matriz de correlaciones. Los coeficientes de correlación pueden ser calculados como se definió en el punto 5.2.

Gráfico N° 8
Mapa de calor



6.3. Determinar las participaciones

Se determinan las participaciones de cada acción en la cartera usando el método propuesto por Markowitz de rendimiento - riesgo. Además, en todo problema de optimización existe un objetivo a optimizar y las restricciones; por ejemplo, el objetivo puede ser minimizar el riesgo y la restricción un nivel de rendimiento.

PyPortfolioOpt implementa los métodos de optimización de carteras. Incluye técnicas clásicas de optimización como la de rendimiento y riesgo (la cuál será usada en el presente trabajo) y la asignación de Black-Litterman, así como desarrollos más recientes como Paridad de Riesgo Jerárquica.

6.3.1. Armar la frontera eficiente

La función `EfficientFrontier` permite definir la frontera eficiente, es decir, el conjunto de carteras que maximiza el rendimiento para determinados niveles de riesgo o minimiza el riesgo para determinados niveles de rendimiento.

El objeto creado con la función contiene varios métodos que corresponden a las diferentes funciones objetivo a optimizar:

- `min_volatility()` : optimiza para obtener la mínima volatilidad, también conocida como cartera de mínima varianza o riesgo.
- `max_sharpe()` : optimiza para obtener el máximo ratio de Sharpe, también conocida como cartera de tangencia.
- `max_quadratic_utility()` : maximiza la utilidad cuadrática, dada una aversión al riesgo.
- `efficient_risk()` : maximiza el rendimiento dado un nivel de riesgo.
- `efficient_return()` : minimiza el riesgo dado un nivel de rendimiento.

```
fe = EfficientFrontier(rendEspl, matrizCovl, weight_bounds=(0, 1))
```

Los parámetros de la función son los rendimientos esperados (vector), la matriz de varianzas y covarianzas y `weight_bounds` que representa el mínimo y máximo peso que puede tomar cada acción en la cartera. En la situación inicial la participación puede ser entre 0% y 100%.

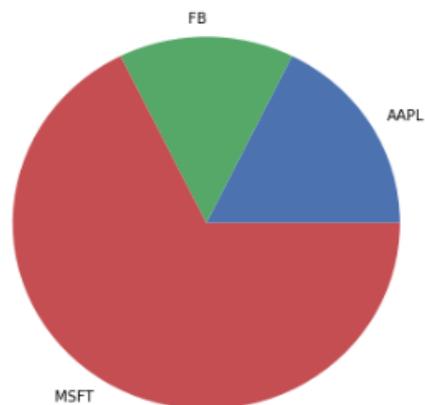
6.3.2. Definir el objetivo de la optimización

6.3.2.1. Minimizar el riesgo

Devuelve los pesos de las acciones para la cartera que minimiza la volatilidad.

```
varianza_minima = fe.min_volatility()
pesos = fe.clean_weights()
pesos
OrderedDict([('AAPL', 0.17703),
             ('FB', 0.1461),
             ('MSFT', 0.67687),
             ('NVDA', 0.0),
             ('TSLA', 0.0)])
```

`clean_weights()` redondea los pesos e iguala a cero los valores muy cercanos a cero.



La cartera de mínimo riesgo estará conformada por MSFT, AAPL y FB en las proporciones del 67,69%, 17,7% y 14,61% respectivamente. El rendimiento esperado, volatilidad y ratio de Sharpe para la cartera optimizada, según el objetivo propuesto es:

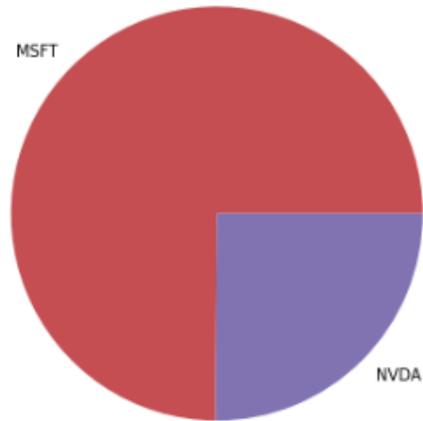
```
Expected annual return: 46.2%
Annual volatility: 20.2%
Sharpe Ratio: 2.19
```

6.3.2.1. Maximizar el ratio de Sharpe

Este método intenta determinar los pesos o las participaciones de las acciones que maximice el rendimiento por unidad de riesgo. Esta cartera es conocida como la cartera de tangencia. Es la cartera donde la línea de mercado de capitales (*capital market line*) es tangente a la frontera eficiente.

Devuelve los pesos de las acciones que maximizan el ratio de Sharpe.

```
max_sharpe = fe.max_sharpe(  
    risk_free_rate=0.02)  
pesos = fe.clean_weights()  
pesos  
  
OrderedDict([('AAPL', 0.0),  
            ('FB', 0.0),  
            ('MSFT', 0.74896),  
            ('NVDA', 0.25104),  
            ('TSLA', 0.0)])
```



Como se puede observar, se aplica una tasa libre de riesgo del 2% anual.

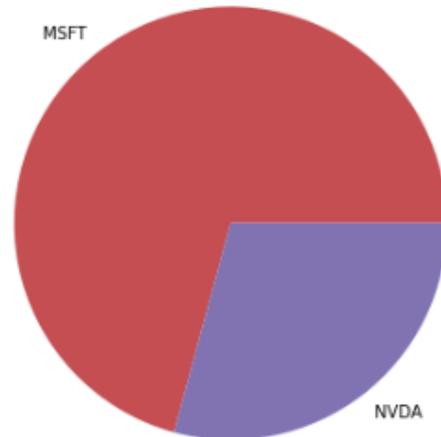
La cartera de máximo ratio de Sharpe estará conformada por MSFT y NVDA en las proporciones del 74,9% y 25,1% respectivamente. El rendimiento esperado, volatilidad y ratio de Sharpe para la cartera optimizada, según el objetivo propuesto es:

```
Expected annual return: 67.1%  
Annual volatility: 24.2%  
Sharpe Ratio: 2.69
```

6.3.2.3. Maximizar el rendimiento dado un nivel de riesgo

Determina las participaciones que maximizan el rendimiento para un nivel dado de riesgo. Es uno de los supuestos básicos de la teoría de Markowitz (Insaciabilidad).

```
fe = EfficientFrontier(  
    rendEsp1, matrizCov1)  
fe.efficient_risk(  
    target_volatility= 0.25)  
pesos = fe.clean_weights()  
pesos  
  
OrderedDict([('AAPL', 0.0),  
            ('FB', 0.0),  
            ('MSFT', 0.70786),  
            ('NVDA', 0.29214),  
            ('TSLA', 0.0)])
```



La cartera que maximiza el rendimiento para un riesgo del 25% estará conformada por MSFT y NVDA en las proporciones del 70,79% y 29,21% respectivamente. El rendimiento esperado, volatilidad y ratio de Sharpe para la cartera optimizada, según el objetivo propuesto es:

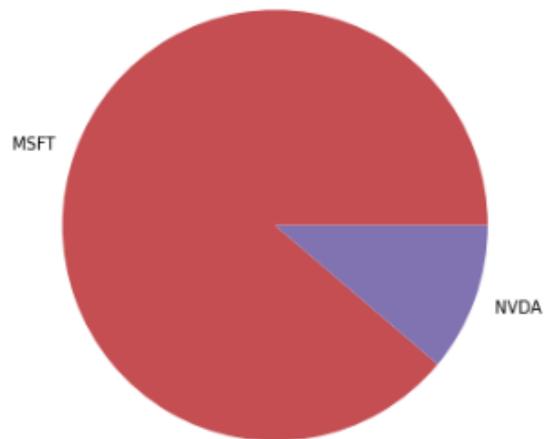
```
Expected annual return: 69.3%  
Annual volatility: 25.0%  
Sharpe Ratio: 2.69
```

3.2.4. Minimizar el riesgo dado un nivel de rendimiento

Determina las participaciones que minimizan el riesgo para un nivel dado de rendimiento. Es otro de los supuestos básicos de la teoría de Markowitz (Aversión al riesgo).

```
fe = EfficientFrontier(
    rendEsp1, matrizCov1)
fe.efficient_return(target_return= 0.6
)
pesos = fe.clean_weights()
pesos

OrderedDict([('AAPL', 0.0),
             ('FB', 0.0),
             ('MSFT', 0.88853),
             ('NVDA', 0.11147),
             ('TSLA', 0.0)])
```



La cartera que minimiza el riesgo para un rendimiento del 60% estará conformada por MSFT y NVDA en las proporciones del 88,85% y 11,15% respectivamente. El rendimiento esperado, volatilidad y ratio de Sharpe para la cartera optimizada, según el objetivo propuesto es:

```
Expected annual return: 60.0%
Annual volatility: 22.0%
Sharpe Ratio: 2.64
```

6.3.3. Agregar objetivos adicionales y restricciones

6.3.3.1. Regularización L2

Este es un concepto del Aprendizaje Automático, sin embargo, acá se usa para incrementar el número de pesos o participaciones distintas de cero y tener una cartera más diversificada.

Se usa como ejemplo la maximización del rendimiento dada una volatilidad del 25%. (participaciones obtenidas: 70% microfofost, 30% nvidia)

Si bien parece que esta cartera cumple con nuestros objetivos, a la mayoría de los *tickers* se les ha asignado peso cero. En efecto, el optimizador se está "sobreajustando" a los datos que ha proporcionado: es más probable que obtenga mejores resultados al aplicar cierto nivel de diversificación. Una forma de lograr esta diversificación es mediante la regularización L2, esencialmente, agregando una penalización a la función objetivo cuando hay ponderaciones (pesos) cercanas a cero.

```
fe = EfficientFrontier(rendEsp1, matrizCov2)
fe.add_objective(objective_functions.L2_reg, gamma=0.1)
fe.efficient_risk(target_volatility= 0.25)
pesos = fe.clean_weights()
pesos

OrderedDict([('AAPL', 0.05617),
             ('FB', 0.0),
             ('MSFT', 0.64926),
             ('NVDA', 0.29458),
             ('TSLA', 0.0)])
```

`gamma` es un parámetro de ajuste, se incrementa en la medida que querramos más pesos distintos de cero (mayor diversificación).

La cartera de mínimo riesgo estará conformada por MSFT, NVDA y AAPL en las proporciones del 64,92%, 29,46% y 5,62% respectivamente. El rendimiento esperado, volatilidad y ratio de Sharpe para la cartera optimizada, según el objetivo propuesto es:

```
Expected annual return: 68.3%
Annual volatility: 25.0%
Sharpe Ratio: 2.65
```

Respecto a la cartera sin regularización L2, se agrega la acción de AAPL en un 5,62% de la inversión.

6.3.3.2. Posiciones Largas / Posiciones Cortas

De forma predeterminada, todos los métodos de optimización de riesgo - rendimiento en *PyPortfolioOpt* toman posiciones largas, sin embargo, se pueden configurar para permitir posiciones cortas cambiando el parámetro `weight_bounds`. Como ya se mencionó, este parámetro representa el mínimo y máximo peso que puede tomar cada acción en la cartera.

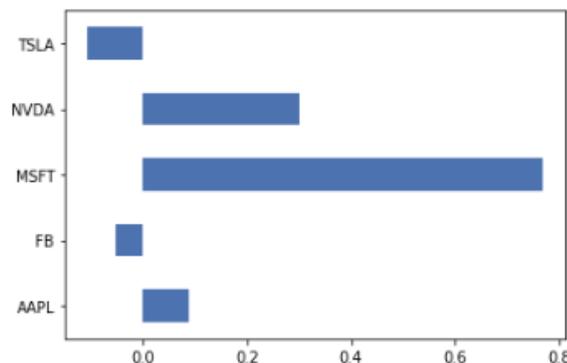
- Posición larga : (0,1) por defecto.
- Posición corta : (-1,1).

Usando el ejemplo anterior:

```
fe = EfficientFrontier(rendEsp1, matrizCov1, weight_bounds=(-1,1))
fe.add_objective(objective_functions.L2_reg, gamma=0.1)
fe.efficient_risk(target_volatility= 0.25)
pesos = fe.clean_weights()
pesos
```

```
OrderedDict([('AAPL', 0.08936),
             ('FB', -0.05145),
             ('MSFT', 0.7677),
             ('NVDA', 0.30083),
             ('TSLA', -0.10644)])
```

Gráfico N° 9
Participaciones



3.3.3. Neutralidad de mercado

Para los objetivos `efficient_risk()` y `efficient_return()`, *PyPortfolioOpt* ofrece una opción para formar una cartera de mercado neutral (es decir, los pesos suman cero). Esto no es posible para la cartera de máximo ratio de Sharpe (cartera de tangencia) y la cartera de mínimo riesgo, porque en esos casos no son invariantes con respecto al apalancamiento. La neutralidad del mercado requiere ponderaciones negativas, es decir, posiciones cortas.

```
fe = EfficientFrontier(rendEsp1, matrizCov1, weight_bounds=(-1,1))
fe.add_objective(objective_functions.L2_reg, gamma=0.1)
fe.efficient_risk(target_volatility= 0.25, market_neutral = True)
```

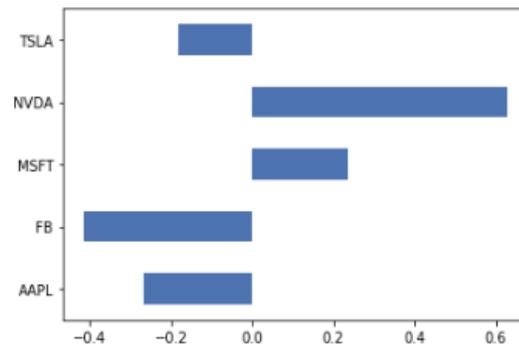
```

pesos = fe.clean_weights()
pesos

OrderedDict([('AAPL', -0.28949),
             ('FB', -0.4289),
             ('MSFT', 0.26823),
             ('NVDA', 0.63001),
             ('TSLA', -0.17985)])

```

Gráfico N° 10
Participaciones



6.3.3.4. Participación en la cartera

En el caso de la máxima participación en la cartera, se busca que cada activo no supere cierta participación dentro de la cartera, para lo cual se modifica el parámetro `weight_bounds`. Por ejemplo, no se podría invertir más del 40% en cada acción:

```

fe = EfficientFrontier(rendEsp1, matrizCov1, weight_bounds=(0, 0.4))
fe.max_sharpe()

OrderedDict([('AAPL', 0.19999999999999999),
             ('FB', 0.0),
             ('MSFT', 0.4),
             ('NVDA', 0.4),
             ('TSLA', 0.0)])

```

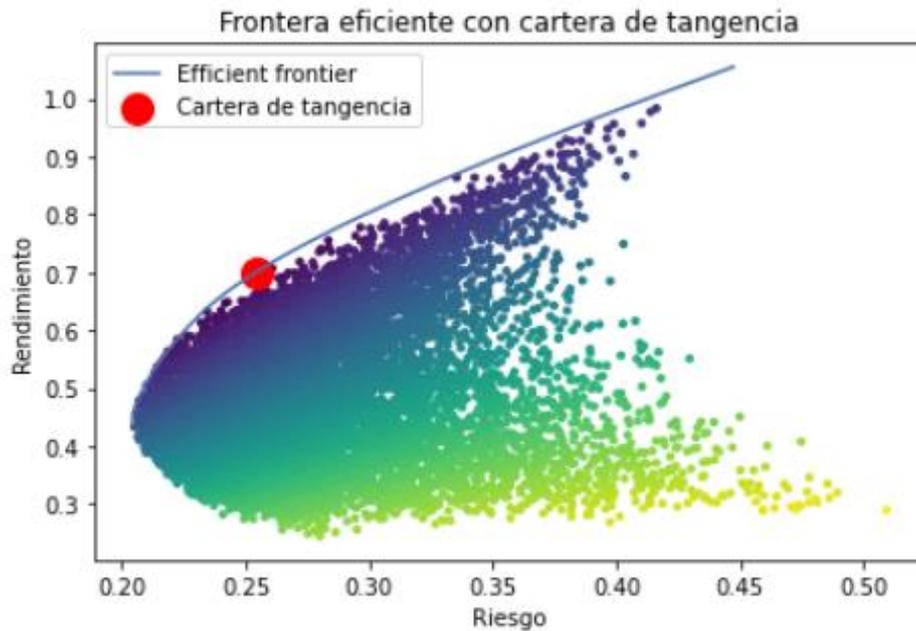
Es posible agregar un sin número de restricciones al modelo usando el método `add_constraint()` y funciones lambda personalizadas. Sin embargo, esto no será tratado en el presente trabajo.

6.4. Graficar carteras

Se graficará la frontera eficiente usando el módulo `plotting` y simulación Montecarlo para generar carteras de manera aleatoria.

En el siguiente ejemplo se graficará la frontera eficiente simulando 100.000 carteras:

Gráfico N° 11



Esta cartera de tangencia es la que maximiza el ratio de Sharpe. (punto 6.3.2.1.).

Además, puede ver el código completo del gráfico y todo el trabajo en el siguiente link:

https://github.com/gustavomachin/PortfolioOptimization_PyPortfolioOpt/blob/master/OptimizacionCarterasPyPortfolioOpt.ipynb

7. Conclusiones

Los resultados para cada objetivo de la optimización, pueden resumirse en el siguiente cuadro:

Tabla N° 5
Resumen objetivos de la optimización

| | Mínimo riesgo | Máximo Sharpe | Máximo rendimiento (riesgo = 25%) | Mínimo riesgo (rendimiento = 60%) |
|------------------------|---------------|---------------|-----------------------------------|-----------------------------------|
| Rendimiento | 46,2% | 67,1% | 69,3% | 60% |
| Riesgo | 20,2% | 24,2% | 25% | 22% |
| Ratio de Sharpe | 2,19 | 2,69 | 2,69 | 2,64 |

El mayor beneficio de usar tecnologías de información a la problemática planteada es la sencillez de aplicación, ya que con unas pocas líneas de código del lenguaje propuesto es posible resolver problemas complejos como la optimización de cartera de acciones. Se recuerda que inclusive con planillas de cálculo se dificulta su resolución. Además, es de acceso gratuito y miles de bibliotecas de código están disponibles para cualquier persona.

En la actualidad es indispensable el manejo de tecnologías de información que faciliten el tratamiento de datos y reduzcan los márgenes de error en los que se incurriría haciendo cálculos, estimaciones o proyecciones manuales. Es por ello que los profesionales de las Ciencias Económicas deben adquirir habilidades y conocimientos para programar y elaborar modelos cuantitativos.

BIBLIOGRAFIA

- Alexander, G., Sharpe W. y Bailey J. (2003). *Fundamentos de inversiones: Teoría y práctica*. Pearson Education.
- Bartolomeo, A. R. y Machín Urbay, G. R., (2013). Optimización de cartera de acciones en condiciones de incertidumbre.

- Elbaum, M. (2004). *Administración de Carteras de Inversión*. Ediciones Macchi.
- Freitas, W. (2021). A curated list of insanely awesome libraries, packages and resources for Quants (Quantitative Finance). Obtenido de: <https://github.com/wilsonfreitas/awesome-quant#data-sources>
- Gujarati, D. y Porter, D. (2009) *Econometría*. McGraw Hill.
- Herrera Lana, E. (2011). *Riesgos en Proyectos de Inversión: Simulación, Pronósticos y Optimización*. Cydhem.
- Hilpisch, Y. (2015). *Python for Finance: Analyze Big Financial Data*. O'Reilly.
- JURCZENKO, E. (2020). *Machine Learning for Asset Management*. Wiley.
- Lewinson, E. (2020). *Python for Finance Cookbook*. Packt Publishing.
- Machín Urbay, G. R., (2022). PortfolioOptimization_PyPortfolioOpt. Obtenido de: https://github.com/gustavomachin/PortfolioOptimization_PyPortfolioOpt
- Martin, R. A., (2021). PyPortfolioOpt: portfolio optimization in Python. *Journal of Open Source Software*, 6(61), 3066. <https://doi.org/10.21105/joss.03066>
- Markowitz, H. (1952). Portfolio Selection. *The Journal of Finance*. 7(1), 77-91.
- Messuti, D., Alvarez, V. y Graffi, H. (1994). *Selección de Inversiones*. Ediciones Macchi.
- Sharpe, W. (1963). A simplified model for portfolio analysis. *Management Science*, 9(2), 277-293.
- Van Horne, J. (1997). *Administración Financiera*. Pearson Education.